The Dissertation Committee for Guan Gong Certifies that this is the approved version of the following dissertation:

Mortality, Education and Bequest

Committee:

Russell Cooper, Supervisor

Li Gan, Co-Supervisor

Stephen Donald

Stephen J. Trejo

Hong Yan

Mortality, Education and Bequest

by

Guan Gong, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2005

Acknowledgements

I am gratefully indebted to many people who have made this dissertation possible. I would like to especially thank the members of my committee Russell Cooper, Li Gan, Stephen Donald, Stephen J. Trejo and Hong Yan for their support, feedback and guidance. I particularly thank Li Gan for his ideas, time, and friendship. Thanks to Vivian Goldman-Leffler for her kindness help. I would also like to thank all my professors at UT for sharing their knowledge and time with me. Finally, thanks to my father Chunyun Gong, my mother Zaihui Tian and my sister Ping Gong for their spiritual and emotional support and unconditional love.

Mortality, Education and Bequest

Publication No._____________

Guan Gong, Ph.D.

The University of Texas at Austin, 2005

Supervisor: Russell Cooper

Co-Supervisor: Li Gan

This dissertation examines the important role mortality risk plays in educational attainment and bequest motives. Chapter 1 investigates the importance of mortality risk in explaining racial difference in education based on a dynamic optimal stopping-point life cycle model. Calibration results show that that more than two-thirds of the empirical difference in education between black and white males can be accounted for by the difference in their mortality differences. Chapter 2 studies interdependence between health and educational attainment. The structural estimation framework fully imposes the restrictions of the existing theoretical hypotheses on the correlation between health and education. The model's estimates imply that an individual's initial health status has a substantial influence on an individual's educational attainment and the expected probability of survival. Policy experiments based on the model's estimates indicate that a health expenditure subsidy conditional on high school attendance would have a larger impact on the educational attainment than a direct college tuition subsidy. Chapter 3 investigates whether subjective expectations about future mortality affect consumption and bequests motives. A dynamic life-cycle model is applied to subjective survival rates and wealth from the panel dataset Asset and Health Dynamics among Oldest Old. The results show that bequest motives are small on average, which indicates that most bequests are involuntary or accidental. Moreover, parameter estimates using subjective mortality risk perform better in predicting out-of-sample wealth levels than estimates using life table mortality risk, suggesting that decisions about consumption and saving are influenced more strongly by individual-level beliefs about mortality risk than by group level mortality risk.

Table of Contents

List of Tables

List of Figures

Chapter 1: Mortality Risk and Educational Attainment of Black and White Men

1.1 INTRODUCTION

Black-white wage and income disparity is a persistent social problem in the United States. A significant body of work has attributed this disparity to forms of discrimination in market places.[1] Anti-discrimination legislation and programs enjoyed early success: blacks reduced the gap with whites during the 1960s and early 1970s. However, the black-white gap stagnated from the 1980s through the early 1990s and then widened in most of the 1990s. In the 1990s, a large body of literature explored whether the black-white gap was a result of factors other than discrimination. Among them, a series of papers found that wage differences between blacks and whites can be explained by differences in their pre-market conditions, especially by differences in their educational attainment. For example, O'Neill (1990) finds that black-white wage differences almost disappear when blacks have the same level of education and Armed Forces Qualifications Test (AFQT) scores as whites. Similar results are found in Maxwell (1994) and Neal and Johnson (1996); Winship and Korenman (1997) and Neal and Johnson (1996) provide convincing evidence that AFQT scores are heavily influenced by the number of years of schooling.

Although there are many interpretations of the factors behind fewer years of schooling for blacks than for whites, in this study we provide a different explanation through mortality risk. Intuitively, education as an investment possesses risk. Although the market education return may be the same for all people, higher mortality risks will

[1] See a survey by Altonji and Blank (1999).

lower the individual return of education and, therefore, might result in fewer years of schooling. Blacks have higher mortality risks than whites, which affords higher risk to reap the wage benefits of schooling. Fewer years of schooling could, then, become blacks' optimal choice.

A large body of studies finds that health and schooling are highly correlated. For example, the life expectancy at birth in England rose from 37.3 to 48.2 years in the nineteenth century and further increased to 60.8 years by 1930. During the same period, the average years of schooling rose from 2.3 years to 9.1 years (see Livi-Bacci [1997] and Matthews, Feinstein and Odling-Smee [1982]). Neoclassical growth literatures interpret the progress on health through the improvement of economic conditions such as gains in per capita income and hence indirectly attribute the gain in health to the improvement of human capital and the relative increase of the amount of schooling. The basic idea of those literatures is that education raises income; a higher income improves nutrition and increase health expenditure, which in turn reduces mortality. Nevertheless, other studies show that the strong relation between health and schooling could reflect the reverse causality, i.e., schooling could be responding to the anticipated amelioration in health. Particularly, some studies argue that health might also be a determinant force behind economic development with its large exogenous component unrelated to scientific knowledge and technological development. For instance, life expectancy in China and Sri Lanka exceeds 70 years, despite these nations having gross national products in 1994 of less than $1,000 per capita (Sen [1999]). Preston (1980, 1996) relates life expectancy changes to income, calorie consumption, and disease, and he concludes that approximately 50% of the changes in life expectancy were due to "structural factors" unrelated to economic development. Fries (1980) states that there is a genetically determined upper limit to life of 85 ± 7 years. Soares (2002) shows that recent reductions in mortality rates across countries were largely independent of improvements in economic conditions.

In this study, we develop a dynamic optimal stopping-point life cycle model in which group-level mortality risk plays an important role in determining individual-level mortality risk, health expenditure, and the amount of schooling. We posit that the mortality risks of the reference group have a negative externality effect on an individual's mortality. In our model, the mortality risks depend not only on health expenditure, but also on the mortality risks of the reference group by which the individual is categorized. Our approach to studying the effect of mortality changes on education is related to the work of Ehrlich and Lui (1991), Kalemli-Ozcan, Ryder, and Weil (2000) and Soares (2002), each of which takes mortality as an exogenous constant to individuals. In contrast, we advance this model to integrate mortality risks into an individual's choices. Although the mortality risks of the reference group for individuals are still taken as exogenous in our framework, the individual-level mortality risks and education are endogenously determined.

The idea of linking the reference group to its effects on mortality risks is nothing new. While genetic traits and lifestyle are usually thought of as the predominant factors explaining health status and mortality, there is a growing consensus that the groups themselves (such as the residential neighborhoods or local community where people live) play an important part in determining people's health. In a related way, a growing literature on social interactions claims that individual outcome is strongly influenced by reference group due to sociological and/or psychological factors (see Manski [1993] and Durlauf [2002]). One bridge linking the reference group to health is the effect of role model or peer group influence in which an individual may desire to conform to the behaviors of older or contemporaneous members of his group and intend to mimic their behaviors. Some of these behaviors are health related, such as smoking, dietary habits, and physical activity, and therefore have detrimental effects on health. Another link between reference group and health is the psychosocial stress caused by diseases or crimes, which may have adverse biological consequences on the individual's health.

Empirical studies finding the relationship between the characteristics of the reference group and the individual's health outcomes are plentiful. Roux et al. (2001) find that living in a disadvantaged neighborhood will increase the incidence of coronary heart disease even after controlling for personal income, education, and occupation. After investigating the influence of individual neighborhood socioeconomic status on mortality, Winkleby and Cubbin (2003) show that a person who lives in a poor neighborhood has a 20% higher death rate than a person who lives in a rich neighborhood after controlling for individual characteristics.

We apply the model to study the impact of differential mortality risks on the educational attainment of black and white men.[2] In particular, we consider an agent at the age of 16, having finished compulsory education, deciding (with his parents) how many additional years of schooling to obtain. The agent faces the probability of death in every period and maximizes the discounted value of the expected utility from consumption and leisure. We assume that a black man has the same utility, discount rate, return to education, and living and working conditions as a white man. The only difference between a black man and a white man is the mortality risk of their reference groups. In this model, both the years of schooling and the life expectancy at the individual level are endogenously determined. The exogenous variable is the reference group's mortality risk.

It is important to point out that we do not claim here that factors such as labor market discrimination, differential opportunities in access to higher education, and parental and occupational preferences do not affect the life prospects of black and white men. What this study shows is that the mortality differences can explain the difference in

[2] We focus our attention only on men. Studying the effect of mortality on education is considerably more difficult for women than for men since women are more likely to face additional choices in leaving the labor force temporarily to have and raise children. Therefore, any meaningful analysis of the effect of mortality risk on education, fertility, and labor force participation for black and white women requires separate treatment.

schooling years obtained when all other factors are the same for both black men and white men.

The model is calibrated to quantify the strength of the effect of mortality risks on schooling. We let the black male population be the reference group for a black man while the white male population is the reference group for a white man. Under a set of reasonable parameter values, our baseline results show that the impact of mortality risk on schooling explains more than two-thirds of the empirical education differences between black and white males. This remains true with a series of sensitivity analyses. Each time we change one of the values of parameters while holding other parameters at their baseline values. We find that although the levels of educational attainment for both blacks and whites deviate from the observed years of schooling, the difference in years of schooling between black men and white men varies little. Since the only difference between blacks and whites is their reference groups' mortality risks, we claim that the observed difference in mortality risks between black men and white men can explain most of their differences in education.

Understanding why blacks have less education than whites has important policy implications. If education is the key reason for future wage differences, public policies designed to reduce the black-white wage gap should concentrate on helping blacks attain more education. If the higher mortality risk of blacks is a major cause of less education, then public policies should put more emphasis on improving access to health care and intervening in the composition of residential neighborhoods, such as making predominantly black neighborhoods safer (since part of the risk may result from living in high-crime neighborhoods).

The rest of this chapter proceeds as follows. Section II introduces a mortality production function and develops a dynamic optimal stopping-point model. In Section III we calibrate the model and explore whether the difference of mortality risks between black males and white males is capable of generating the observed educational difference. A brief conclusion is given in Section IV.

1.2 THE MODEL

In this section, we develop a dynamic optimal stopping-point model to analyze the effects of mortality risks on years of schooling. We start with the mortality risk production function since it explicitly illustrates the channel from mortality to schooling.

We assume that an individual's production function of mortality risk is:

$$m(t) = \mu \hat{m}(t) e^{-\beta d(t)}, \tag{1.1}$$

where m is the hazard rate of the individual, $\hat{m}$ is the hazard rate of the reference group, and d is the health expenditure. The term "health expenditure" includes all expenditures that may affect an agent's mortality, such as time and money spent on health clubs, appropriate nutrition, medical insurance, and other expenses related to health care. Spending on smoking can also be included as a negative expenditure. Health expenditure as an input into the production function of mortality has been broadly used in the literature of health economics since Grossman (1972a, 1972b). Individual-specific health characteristics, such as genetic traits and illness, are captured by a positive parameter μ. The exponential specification in (1.1) implies that health expenditure has a decreasing marginal effect on health, with β being the percentage gain in mortality reduction from one unit of health expenditure.

The negative externality of the reference group's mortality on individual health, as argued in Section 1.1, has been substantiated in growing studies. Although the mortality risk of the reference group is taken as exogenous to the individual in the model, the current model does not offer any guidance as to how the reference group is selected. For example, a black male may choose the general black male population as his reference group. Alternatively, he may view a smaller group of people whom he is familiar with, such as his family and friends, as his reference group. It is also possible for a black male living in a suburban white neighborhood to view the white male population as his reference group. In other words, identifying the reference group may be subjective.

It is worth noting that (1.1) assumes that an agent's mortality is only affected by his reference group's mortality and his health expenditure. The agent's education affects his mortality only by health related expenditure. However, previous literature shows that a better-educated agent can be more effective in using the money he spends on reducing mortality (or productive efficiency). In addition, since a better educated agent may have more knowledge of the adverse effects of some activities (i.e., smoking, bad diet, etc.) and the positive effects of other activities (i.e., exercise, appropriate diet, etc.), he is more likely to allocate resources to improve his health (or allocative efficiency).[3] Modeling these two efficiencies in the current framework is beyond the scope of this study.

With the mortality risk production function in (1.1), the survival rate for the individual is given by:

$$p(t) = \exp\left\{-\mu\int_0^t \hat{m}(t)\exp[-\beta d(t)]\,dt\right\}. \tag{1.2}$$

Equation (1.2) implies that the level of survival rate for an individual is a positive function of his current and past health expenditure.

Now we turn to the dynamic optimal stopping-point model. We consider an individual who is 16 years old. After finishing his compulsory years of schooling, he (with his parents) chooses how many additional years of schooling he will undertake. Let the instantaneous utility at time *t* be $u(c(t), l(t))$, where $c(t)$ is consumption and $l(t)$ is leisure (the labor supply is $1-l(t)$). The function $u(\cdot,\cdot)$ is assumed to be strictly concave, increasing in each argument, twice continuously differentiable. The lifetime utility maximization problem is, in the formulation of Yaari (1965):

[3] For a survey on productive efficiency, see Grossman (2000); the survey on allocative efficiency can be found in Kenkel (2000).

$$\max_{S,\, d(t), c(t), l(t)} \int_0^S u(c(t), \bar{l})p(t)e^{-\theta t}dt + \int_S^T u(c(t), l(t))p(t)e^{-\theta t}dt + \int_T^N u(c(t), 1)p(t)e^{-\theta t}dt \quad (1.3)$$

where choice variable S is the amount of additional schooling after 9 years of compulsory schooling, θ is the time discount rate, T is the time of retirement, and N is the maximum longevity. In this model, we let the retirement age and maximum longevity be exogenous. We assume that the individual retires at age 65, thus $T = 49$ (i.e., age 65 minus the initial age 16). And we let the maximum age to which the individual could survive be 110, thus $N = 94$ (i.e. age 110 minus the initial age 16). The only uncertainty that the agent faces at any future date comes from the possibility of death.

The lifetime utility in (1.3) consists of three parts, representing three stages of the individual's life cycle. The first part in (1.3) is the expected utility from schooling. At the schooling stage, we assume that schooling is structured such that leisure from schooling in each period is a constant, $l(t) = \bar{l}$ for $t<S$. The individual chooses additional years of schooling, S, and a consumption profile at this stage. The life cycle model in (1.3) assumes irreversibility: if an individual has started to work, he cannot come back to school again at later time in his life cycle. The second part in (1.3) is the expected utility from working. At this stage, the individual chooses a profile of consumption and leisure. At time T, the agent retires from work. The third part in (1.3) describes the expected utility from retirement. At this stage, the agent only chooses a consumption profile. His leisure after retirement is 1.

Corresponding to the life cycle utility function in (1.3), the agent's wealth (or asset) accumulation equation is divided into three parts:

$$\begin{aligned}
t \in [0,S): \quad & \dot{A}(t) = rA(t) - c(t) - \xi(1-\bar{l})wh(t) - d(t), \\
& \dot{h}(t) = g(t)h(t); \\
t \in [S,T): \quad & \dot{A}(t) = rA(t) + wh(t)(1-l(t)) - c(t) - d(t), \\
& \dot{h}(t) = 0; \\
t \in [T,N]: \quad & \dot{A}(t) = rA(t) - c(t) - d(t); \\
& A(0) = A(N) = 0, \ \ h(0) \text{ given,}
\end{aligned} \qquad (1.4)$$

where $A(t)$ is wealth at time t, $h(t)$ is the human capital at time t, and w is wage rate per unit of human capital. The market interest rate r is assumed to be constant. At the first stage, the individual accumulates human capital with the rate $g(t)$ at each t. Function $g(\cdot)$ is increasing and concave in the amount of schooling. Following Bils and Klenow (2000), we assume that the cost of education (including tuition, room and board) increases with the level of education. The parameter $\xi(>0)$ is the ratio of schooling cost to the opportunity cost of student time. At the second stage, the agent goes to work and earns the labor income, defined as per unit of labor wage ($wh(t)$) multiplied by his labor supply ($1-l(t)$). For convenience, we assume that there is no accumulation or depreciation of human capital at this stage. At the third stage, the agent retires and consumes the wealth he accumulated when he worked. The initial and end wealth are assumed to be zero, and the initial human capital is given.

The first-order conditions yield the differential equations for consumption:[4]

$$t \in [0,S): \quad \frac{\dot{c}(t)}{c(t)} = -\frac{u_c\left(c(t),\bar{l}\right)}{c(t)u_{cc}\left(c(t),\bar{l}\right)}\left(r-\theta-m(t)\right) \tag{1.5a}$$

$$t \in [S,T): \quad \dot{c}(t) = \frac{u_{cl}\left(c(t),l(t)\right)u_l\left(c(t),l(t)\right) - u_{ll}\left(c(t),l(t)\right)u_c\left(c(t),l(t)\right)}{u_{cc}\left(c(t),l(t)\right)u_{ll}\left(c(t),l(t)\right) - \left[u_{cl}\left(c(t),l(t)\right)\right]^2}\left(r-\theta-m(t)\right) \tag{1.5b}$$

$$t \in [T,N]: \quad \frac{\dot{c}(t)}{c(t)} = -\frac{u_c\left(c(t),1\right)}{c(t)u_{cc}\left(c(t),1\right)}\left(r-\theta-m(t)\right). \tag{1.5c}$$

Equations (1.5a)–(1.5c) are the ordinary Euler equations respectively corresponding to different stages. These three equations describe necessary conditions that have to be satisfied on any optimal path. At any $t \in [S,T)$, the optimal consumption and leisure make the marginal rate of substitution equal to the marginal rate of transformation:

4 For solving a dynamic optimization problem with switches in the state equations, see Kamien and Schwartz (1991).

$$\frac{u_l(c(t),l(t))}{u_c(c(t),l(t))}=wh(t). \tag{1.6}$$

At the times S and T, there are jumps in consumption and leisure. The consumption and leisure at these two points satisfy the conditions:

$$u_c\left(c\left(S^-\right),\bar{l}\right)=u_c\left(c\left(S^+\right),l\left(S^+\right)\right), \quad \text{and} \quad u_c\left(c\left(T^-\right),l\left(T^-\right)\right)=u_c\left(c\left(T^+\right),1\right), \tag{1.7}$$

where S^- is defined as $t < S$ and $t \to S$, while S^+ is defined as $t > S$ and $t \to S$. The variables T^- and T^+ are analogues to S^- and S^+. Equation (1.7) says that the optimal consumption and leisure will make the marginal utility of consumption be the same at the time when the agent switches from one stage to another stage (i.e., from schooling to working and from working to retirement).

The optimal health expenditure satisfies

$$\beta\int_t^N u(c(v),l(v))p(v)m(v)e^{-\theta(v-t)}dv = u_c(c(t),l(t))p(t), \tag{1.8}$$

where the left-hand side (divided by the right-hand side) is the change (in monetary unit) of the present value of utility from increases in current and future survival rates caused by health expenditure. Therefore, equation (1.8) implies that the necessary condition for optimal health expenditure equates the marginal gain from an extra unit of health expenditure to its marginal cost, which is 1.

Finally, the necessary condition for the optimal amount of schooling is:

$$\begin{aligned}\frac{u\left(c\left(S^-\right),\bar{l}\right)-u\left(c\left(S^+\right),l\left(S^+\right)\right)}{u_c\left(c\left(S^+\right),l\left(S^+\right)\right)} &= \left(c\left(S^-\right)-c\left(S^+\right)\right) \\ &+wh\,(S)\left[\xi\left(1-\bar{l}\right)+1-l\left(S^+\right)-g(S)\int_S^T\left(1-l(t)\right)e^{-r(t-S)}dt\right].\end{aligned} \tag{1.9}$$

Equation (1.9) implies that marginal gains are equal to marginal costs from an extra year of schooling. The marginal gains include gain in utility, $\left[u\left(c\left(S^-\right),\bar{l}\right)-u\left(c\left(S^+\right),l\left(S^+\right)\right)\right]/u_c$, and gain in future earnings discounted to present, $-wh(s)g(s)\int_S^T (1-l(t))\exp(-r(t-S))dt$. The marginal costs of schooling include consumption $\left(c(S^-)-c(S^+)\right)$, tuition $\xi\left(1-\bar{l}\right)wh\left(S\right)$, and opportunity cost from forgoing working $\left(1-l\left(S^+\right)\right)wh\left(S\right)$ by staying in school. The gap between the utility from attending schooling and that from going to work enters because of the jumps of consumption and leisure at the time of the switch in stages. The same reason applies to the gap of consumptions in equation (1.9).

The individual optimal amount of schooling and hazard rate are not explicit functions of the model's parameters and the mortality risks of the reference group. In the next section, we apply the model to the calibration method and explore to what extent the difference in educational attainment between black and white males can be attributed to the difference in their mortality risks. It is important to recognize that when studying the differential mortality risks between black and white men, it is necessary that the model work with age-varying mortality risks since black and white men have different mortality risk patterns over their life cycles.

1.3 MORTALITY RISK AND EDUCATIONAL ATTAINMENT OF BLACK AND WHITE MEN

In this section, we apply our earlier model to study the main objective of the investigation: to find out to what extent the differences in education between black men and white men can be explained by their difference in mortality risks.

It is well known that mortality risks are different for black and white men. In the 1979–1981 U.S. decennial life tables, the life expectancy (conditional on surviving to age 16) is 66.2 years for a black male and 72.1 years for a white male. Relative average mortality risks vary for different age groups. For example, for people ages 21–30, the average yearly mortality risk is .311% for black men, which is 75% higher than the mortality risk of white men, or .178%. For people ages 31–40, the average yearly mortality risk for black men is .440%, which is 159% higher than the mortality risk for white men, or .167%.

In our framework, since an agent's reference group is subjective, it is difficult for researchers to determine an agent's exact reference group.[5] However, in some cases researchers should be able to determine what an agent's reference group is most likely to be. For example, given that blacks are very likely to live in neighborhoods with few whites (Massey and Denton, 1989), researchers should be confident that the reference group of a black male is likely to consist of a majority population of black males; similarly, a white male's reference group should have a preponderance of white males.[6]

We use the U.S. decennial life tables in 1979–1981 to represent the mortality risks that people in an age group observe when they make their decisions about years of schooling. The years of schooling are based on 1990 census data. The average years of schooling for black men ages 26–36 who were in the labor force in 1990 were 12.74 years, while the same group of white men averaged 13.50 years. We concentrate on men ages 26–36 in 1990 for two reasons. First, people in this age group have already finished their education. Second, since people with less education have higher mortality rates, selecting a relatively young group will minimize that sample-selection problem.

The rest of this section includes three parts. In the first part, we set the baseline parameter values for the model to calibrate the optimal years of schooling. In the second part, we report the results from calibration compared with the observed years of

---

5 Some authors argue that groups can be endogenously determined. For example, Fernandez and Rogerson (1996) show that individuals endogenously select themselves into different communities or groups according to income.

6 This paper does not investigate why exogenous difference in mortality between blacks and whites exists. One possibility is that rampant discrimination in the labor force before the civil rights movement in the 1960s caused a lower return to education for blacks than for whites. As a consequence, blacks took less education and spent less in health care than whites, resulting in a higher mortality risk than that of whites.

schooling. In the third part, we conduct a series of sensitivity analyses by letting parameters deviate from baseline parameter values.

1.3.1 Baseline Parameters and Utility Functional Forms

Applying the model to explore the effect of mortality differences between black and white men on their education differences requires parameterized functional forms for mortality risk, utility, and human capital. We first begin by calibrating the production function for mortality risk.

In equation (1.1), the parameter β is the percentage reduction in the average mortality risk from one unit of health expenditure. According to Jones (2002), the life expectancy in the United States was 66.6 years in 1960 and 73.9 years in 1997. Thus, the average yearly mortality risk was approximately lowered from 1/66.6 in 1960 to 1/73.9 in 1990, a reduction of about 9.88%. In the meantime, the U.S. per capita health expenditure rose from \$504.60 in 1960 to \$2,127 in 1997. Therefore, a \$10,000 increase in health expenditure will, on average, reduce mortality risk by: 9.88%*10,000/(2,727 – 504.60) = 0.445. We take the value of parameter β as 0.445, meaning that a \$10,000 health expenditure will reduce mortality risk by 44.5%. Note that the current calculation of β assumes that the group mortality $\hat{m}(t)$ is constant over time. If we let $\hat{m}(t)$ be a function of health expenditure such that $\partial \hat{m}(t) / \partial d(t) > 0$, the value of β is overestimated. In the sensitivity analysis in Section IIIC, we discuss how the outcomes of the model vary when β varies.

To calibrate the value of parameter μ, we rewrite equation (1.1) as the following log form:

$$\ln m(t) = \ln \mu + \ln \hat{m}(t) - \beta\, d(t)\,. \tag{1.10}$$

The value of μ can be calculated by taking the mean on the natural log of mortality risks across individuals in the reference group. Since $\hat{m}(t)$ is the group mortality, i.e., $E[m(t)] = \hat{m}(t)$, we must have $E[\ln m(t)] = \ln \hat{m}(t) - c$, where $c > 0$ (by Jensen's inequality). Since no guidance is offered in the literature on the value of c, we calibrate the baseline value μ by assuming that $c = 0$. In particular, when $c = 0$, the ratio $\ln(\mu)/\beta$ matches the mean health expenditure in the reference group. The baseline value of μ is calculated using the U.S. health expenditure (\$2,166.50 or 12% of GDP) in 1990.[7] In this case μ =1.101. That is to say that if the individual's health expenditure is zero, his mortality risk is around 10% higher than that of his reference group. If the constant $c > 0$, the parameter μ is smaller. Therefore, the baseline parameter value μ is larger than the real parameter value μ. We discuss how the outcomes of the model vary if μ changes in the sensitivity analysis in Section IIIC.

Then, we come to the utility function, which is given by:

$$u(c,l) = \frac{\left(c^{\alpha} l^{1-\alpha}\right)^{1-\sigma} - 1}{1-\sigma}, \tag{1.11}$$

where the relative risk aversion parameter, σ, is set to 2, and the consumption share in utility, α, is set to equal 0.33, as in Backus, Kehoe and Kydland (1994).

Based on the utility function in (1.11), the first-order conditions (1.5) –(1.7) say that for any $t \in [0, S)$, the consumption is:

$$c(t) = c(0) p(t)^{\frac{1}{1-\alpha(1-\sigma)}} e^{\frac{(r-\theta)}{1-\alpha(1-\sigma)} t} \quad ; \tag{1.12}$$

at the switching point from staying in school to going to work, S, the consumption satisfies:

$$c(S^{+}) = \left(\frac{1-\alpha}{\alpha \bar{l} w h(S)}\right)^{\frac{(1-\alpha)(1-\sigma)}{\sigma}} c(S^{-})^{\frac{1-\alpha(1-\sigma)}{\sigma}} ; \tag{1.13}$$

---

7 U.S. health expenditure data are from the Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group, National health expenditures, 2001. Internet address: www.cms.hhs.gov/statistics/nhe.

for $t \in [S,T)$, the leisure $l(t)$ and consumption $c(t)$ are given by equations (1.14) and (1.15):

$$l(t) = \frac{1-\alpha}{\alpha w h(S)} c(t), \tag{1.14}$$

$$c(t) = c(0)^{\frac{1-\alpha(1-\sigma)}{\sigma}} p(t)^{\frac{1}{\sigma}} \left(\frac{1-\alpha}{\alpha \bar{l} w h(S)}\right)^{\frac{(1-\alpha)(1-\sigma)}{\sigma}} e^{\frac{(r-\theta)}{\sigma}t}; \tag{1.15}$$

at the switching point from working to retirement, T, the consumption satisfies:

$$c(T^+) = \left(\frac{1-\alpha}{\alpha w h_S}\right)^{\frac{(1-\alpha)(1-\sigma)}{\alpha(1-\sigma)-1}} c(T^-)^{\frac{\sigma}{1-\alpha(1-\sigma)}}; \tag{1.16}$$

and for $t \in [T,N]$, the consumption is:

$$c(t) = c(0) p(t)^{\frac{1}{1-\alpha(1-\sigma)}} \bar{l}^{\frac{(1-\alpha)(1-\sigma)}{\alpha(1-\sigma)-1}} e^{\frac{(r-\theta)}{1-\alpha(1-\sigma)}t}. \tag{1.17}$$

The optimal amount of schooling satisfies the equation:

$$\frac{1-\alpha(1-\sigma)}{\alpha(1-\sigma)} c(S^-) - \frac{\sigma}{\alpha(1-\sigma)} c(S^+) = w h(S)\left[1 + \xi(1-\bar{l}) - g(S)\int_S^T (1-l(t)) e^{-r(t-S)} dt\right]. \tag{1.18}$$

Finally, following Bils and Klenow (2000), we let $g(t) = \eta(t+9)^{-\varphi}$. The term $(t+9)$ reflects the fact that the agent has finished 9 years of compulsory schooling. The human-capital accumulation is given by:

$$h(t) = \exp\left(\frac{\eta}{1-\varphi}(t+9)^{1-\varphi}\right), \tag{1.19}$$

where $\eta = 0.32$, and $\varphi = 0.58$, as in Bils and Klenow (2000). In this setup, the marginal return of schooling is decreasing. At the given parameter values, the return of an additional year of education is about 9% if a person has just finished nine years of compulsory schooling.

Other parameters used in the calibration are chosen as the following values: the time discount factor is $\theta = 0.032$; interest rate r is 0.04; and the parameter governing the education cost ξ is 0.5, as in Bils and Klenow (2000). The wage rate per hour for one unit of human capital is \$1.47, at which a person with nine years of compulsory schooling will earn \$10 per hour.[8] Since there is no guidance in the literature about the value of leisure during schooling, we let $\bar{l} = 0.4$, i.e., when a person is in school, he uses 60% of his expendable time on studying.

1.3.2 Results

Given the baseline values for various parameters, we can then obtain the optimal quantity of schooling, paths for consumption and mortality, and optimal levels of health expenditures based on equations (1.12) –(1.18). However, solving this optimization problem with mortality risk turns out to be numerically challenging. We restrict the analysis to a time independent health expenditure, i.e., $d(t) = d$. This assumption greatly simplifies the solution.[9]

The results from the baseline parameters are denoted as baseline results. Before we present our baseline results, a simple normalization is worth mentioning here. In our analysis, the unit of time is one year, denoted as 1. All reported parameter values in our study (in Tables 1.1 and 1.2 and in Figures 1.1 and 1.2) correspond to this. In order to discuss our result in more intuitive dollar values, we assume that the total hours that an agent can allocate between leisure and work in a year is 5,000, reflecting about 13.7 hours per day.[10] The upper panel of Table 1.1 lists the baseline parameter values and the lower panel reports the baseline results. The baseline results show that the optimal years

---

[8] Based on Census 1990, the average hourly rate for men with only nine years of schooling is \$9.95.

[9] When the health expenditure d is constant across ages, the necessary condition for the optimal health expenditure is given by:

$$\beta r \int_0^N u(c(v), l(v)) p(v) \ln[p(v)] e^{-\theta v} dv = u_c(c(0), \bar{l})(e^{-rN} - 1)$$

[10] If one assumes that the total hours per year are 4,000, then all the dollar values reported in Table 1 and Table 2 will be proportionally lower. However, the schooling years are not affected by the total hours per year assumed in the model.

of schooling is 12.6 years for black men and 13.12 years for white men. Compared to the observed 12.74 years of schooling for black men and 13.50 years for white men, the predicted schooling years are a little lower and the predicted gap in schooling is 68.4% of the observed gap. The average hourly wage rate at the predicted years of schooling is $13.38 for blacks and $13.89 for whites.

The predicted health expenditures from the model are $1,584 for blacks and $1,802 for whites. This suggests that white men spend about 20% more than black men in health expenditures. Given that the predicted schooling years are lower than the observed schooling years, it is not surprising that predicted health expenditures from the model are lower than the U.S. per capita health expenditure in 1990 ($2,167).

Figure 1.1 illustrates the predicted lifetime trajectories of consumption, income, wealth and leisure. In Figure 1.1a where the consumption trajectories are shown, one interesting observation is the large drop in consumption level at the time of retirement. Based on equation (1.7), the marginal utility just before retirement should equal the marginal utility just after retirement. Since leisure and consumption are substitutable in the given utility function, an increase in leisure due to retirement is compensated by a lower consumption of goods. In Table 1.1, blacks spend an average $12,500 in consumption per year, while whites on average consume $13,471 per year. Whites consume 7.77% more than blacks.

Figure 1.1b shows lifetime trajectories for net incomes, defined as the labor income minus the sum of health expenditures and the cost of schooling. At the working stage, blacks' average labor income is $18,850, while whites' average income is $20,059. Whites' labor incomes are 6.4% higher than blacks' labor incomes. For the reason of simplicity, our model does not include returns of experience in the accumulation of human capital. In our setup, wages for both blacks and whites do not increase after they finish school.

The wealth trajectories in Figure 1.1c show a familiar life cycle pattern: both the black and the white agent borrow to finance their education, save when they work, and dissave after they retire. The black agent's wealth level is lower than the white agent's during most of the life span. The only period when the black agent's wealth exceeds the white agent's is the period immediately after schooling, since the black agent starts to work earlier than the white agent. The maximum wealth for both blacks and whites occurs at age 65 when they are about to retire. The maximum wealth level is $307.1K for blacks and $360.1K for whites. The lifetime mean wealth level is $97,890 for blacks and $116,970 for whites. White men have 19.5% more wealth than black men.

In the trajectories of leisure in Figure 1.1d, schooling requires more studying hours (or less leisure) than working. During the second stage, when people work, the labor supply of black men is slightly lower than the labor supply of white men, indicating that black men not only have less education, but they also work less. During the working stage, the labor supply for the black agent is 2,272 hours per year, while the labor supply for the white agent is 2,314 hours.

In summary, in this model, blacks and whites are the same except in the mortality risks of the reference group by which they are categorized. Therefore, all the differences in economic outcomes, including consumption, income, wealth, and labor supply, are attributed to the differences in mortality risks from the reference groups. More than two-thirds of the black-white educational difference can be explained by their difference in mortality risk.

1.3.3 Sensitivity Analysis

In the previous subsection, we show that when parameters are given their baseline values, the predicted schooling difference is over two-thirds of the observed difference between black men and white men. In this subsection, we study whether the baseline results hold beyond the particular set of parameter values.

The sensitivity analysis is conducted according to the following procedure. We let one parameter vary at a time while holding other parameters constant at their baseline values. For any new set of parameter values, we re-optimize the whole life-cycle model to obtain optimal years of schooling for blacks and whites. For each parameter, we must determine a parameter interval in which we may conduct a sensitivity analysis. Selecting the parameter interval involves two steps. First, we search the boundary parameter value. When the parameter is beyond the boundary value, the additional years of schooling for blacks are negative (i.e., the total years of schooling are fewer than the minimum nine years of schooling assumed in the study), or no solution can be found. Second, we let the middle point of the interval be the baseline parameter value, and we let one end of the interval be the boundary parameter value we just selected in the first step. Obviously, the interval is determined after one end point and the middle point of the interval are chosen. For example, for the time discount rate θ, the baseline parameter is $\theta = 0.032$. First, we find out that when $\theta > 0.034$, optimal years of schooling for blacks would be negative. Second, when we let $\theta = 0.034$ be the upper boundary of the parameter interval and let $\theta = 0.032$ be the middle point, the lower boundary of the interval is then $\theta = 0.030$. Thus, the interval to conduct sensitivity analysis for the time discount rate is [0.030, 0.034]. This interval is then divided into 20 equally spaced sub-intervals. There are 21 end points of these 20 sub-intervals. We let θ be each of these 21 end points. For each different θ, we obtain optimal schooling years and health expenditures. We obtain 21 sets of schooling years and health expenditures for both black men and white men, one of which is the baseline result.

With these 21 sets of schooling years and health expenditures, we calculate the mean differences and their standard errors in schooling years and in health expenditures between black men and white men. This process repeats for other parameters: leisure in school $\bar{l}$, cost of education parameter ξ, risk averse parameter σ, mortality production parameter β and μ, and the interest rate r. The returns of education are calculated at nine years of schooling. From equation (1.19), there are two parameters that determine the return of education. For simplicity, we let only the parameter φ in (1.19) change to obtain the parameter interval for the return of education.

From Table 1.2, we see that the mean differences in years of schooling under various experiments are very similar. The observed black-white difference is 0.76 years. When we let the time discount rate θ vary between 0.030 and 0.034, the mean difference in years of schooling is 0.59, which is a little higher than two-thirds of the observed difference. In fact, the mean differences range from 0.537 to 0.646 when all parameters except the interest rate vary in their parameter intervals. When the interest rate varies in its parameter interval, the mean difference in schooling years is 0.890, which is larger than the observed difference in schooling years. We conclude that the impact of mortality risk on schooling explains more than two-thirds of the empirical education difference in schooling years between black and white men.

The baseline parameter values for the mortality production function in (1.1) are β = 0.445 and μ = 1.101. In Section IIIA, we show that the baseline values likely overestimate the actual parameter values. Here we discuss the outcomes of the model if either of the two parameters have lower values. We consider lowering the parameter μ. For example, if μ is lowered by 20%, i.e., μ = 0.9, the schooling years are 13.10 for blacks and 13.62 for whites. The difference between blacks and whites remain the same as the baseline case. In fact, if we let $\mu \in [0.701, 1.501]$, the average difference in schooling years between blacks and whites is 0.565 with a standard deviation of 0.185. The difference in schooling years is quite robust to the value of μ. However, the difference in schooling years is more sensitive to the parameter β. For example, if we lower β by 20% (while other parameters are at their baseline values) i.e. β = 0.356, the schooling years are 12.463 for whites and 12.077 for blacks. Although the difference in schooling years is reduced to 0.386 years, it represents a significant portion (50%) of the observed difference in schooling years.

In addition to the differences in schooling years, Table 1.2 also lists mean differences in health expenditures and their standard errors for all other parameters. For example, when the time discount rate varies in the interval [0.030, 0.034], the health expenditures for blacks vary from \$826 (when θ = 0.030) to \$1,698 (when θ = 0.034). The whites' health expenditure varies from \$1274 (when θ = 0.030) to \$2,119 (when θ = 0.034). The mean difference in health expenditures, when the time discount varies, is \$565 with a standard error of \$152. In fact, a different set of parameters produces a different set of health expenditures for both blacks and whites.

Figure 1.2a-h illustrates the schooling years of blacks and whites when each of the parameters varies in its parameter interval. For example, Figure 1.2a shows schooling years when the time discount rate θ varies in its parameter interval, [0.030, 0.034], while other parameters are held at their baseline values. The schooling years for whites lie above the schooling years for blacks. Although the level change of schooling years is rather large, from 10.66 years to 13.86 years for blacks and from 11.25 years to 14.62 years for whites, as the time discount rate increases from 0.030 to 0.034, the difference in years of schooling stays roughly the same. The standard error of the average difference in schooling years between whites and blacks is 0.038, only about 6% of its mean. Therefore, when the time discount rate varies in its parameter interval, the level of schooling years is no longer consistent with the observed years of schooling. However, the black-white difference in schooling years from our model is consistent with the observed difference.

Similar patterns repeat for four other parameters: leisure in school (Figure 1.2b), mortality function parameter μ (Figure 1.2c), the return of education (Figure 1.2d), and the cost of education (Figure 1.2e). When one of these four parameters varies in its respective parameter intervals, levels of schooling years vary greatly; however, the mean differences (with relatively small standard errors) in black-white schooling years match with the observed difference. Therefore, the result showing that the difference in mortality risks can explain much of the black-white difference in schooling years is robust for these four parameters.

For the remaining three parameters, the risk averse parameter σ, the mortality production parameter β, and the interest rate r, schooling years for whites always lie above those for blacks, indicating that whites always complete more schooling years than blacks. In addition, the mean differences match with observed differences in schooling years when each of these parameters varies in its respective parameter intervals. However, these mean differences in schooling years have a larger variation. For the risk averse parameter σ, the difference in black-white schooling years varies from 0.115 (σ = 2.26) to 0.843 (σ = 1.74). The average difference is 0.537 years with a standard error of 0.203. For the mortality production parameter β, the difference in black-white schooling years varies from 0.103 (β = 0.245) to 0.815 (β = 0.645); the average difference is 0.638 with a standard error of 0.253. Finally, when the interest varies from 0.026 to 0.054, the average difference is 0.890 with a standard error of 0.733. Since the mean differences in schooling years from our model are consistently around two-thirds of the observed difference in schooling years, we claim that the difference in schooling years for black and white men can be substantially explained by the mortality risks. However, such a claim is less robust for three out of the eight parameters discussed in the study.

Finally, from Figure 1.2a – Figure 1.2h, one can find out how choices in schooling years change when one of the parameters changes. The figures are rather intuitive. When the leisure in school is higher, staying in school becomes more appealing and the years of schooling increase (Figure 1.2b). In Figure 1.2c, when the mortality production parameter varies, the marginal gain from health expenditures increases. Therefore, it is beneficial to have more education in order to afford better health expenditures. Similar reasoning applies to Figure 1.2g. A higher return of education raises years of schooling (Figure 1.2d), while a higher cost of education lowers years of schooling (Figure 1.2e). In Figure 1.2h, a higher interest rate lowers years of schooling since it raises the opportunity cost of schooling. The intuition in other figures is only slightly more complicated. In Figure 1.2a, a higher discount rate lowers years of schooling since current utility is valued higher. In Figure 1.2f, a more risk averse person has lower years of schooling since he has a higher tendency to avoid risky investment of education.

1.4 Conclusion

Tremendous resources have been devoted to reduce the black-white gap. This study investigates to what extent the difference in educational attainment between black and white men can be explained by the differences in their mortality risks. We develop a dynamic life-cycle model with an optimal stopping-point in which group-level mortality risk plays an important role in determining individual-level health expenditure, mortality risk, and amount of schooling. In the model, an agent's mortality is a function of his own health expenditure and his reference group's mortality risks. In such a framework, both the agent's years of schooling and mortality risks are endogenous while the reference group's mortality risks are exogenous.

We let the black male population be the reference group for a representative black male and the white male population be the reference group for a representative white male. The resulting years of schooling for black and white men are then compared with observed schooling for black and white men, respectively.

We calibrate the model by finding a set of baseline parameter values such that optimal schooling years match a large part of the observed years of schooling for both black men (12.74 years) and white men (13.50 years). The optimal health expenditures are $1,584 per year for a black male and $1,802 per year for a white male, meaning that blacks spend about 12% less in health expenditure than whites. We then conduct various sensitivity analyses by locally varying parameters. We find that although levels of schooling years are sensitive to various parameter values, the difference in schooling years between blacks and whites is relatively robust in various parameter values. We conclude that the mortality difference between blacks and whites is capable of explaining their difference in educational attainment.

Table 1.1: Baseline Parameter Values and Results

Parameter description and notation		Values	
Mortality production function:			
	parameter β	0.445	
	parameter μ	1.101	
Utility function:			
	relative risk averse σ	2.0	
	share of consumption α	.33	
Human capital function:			
	parameter φ	0.58	
	parameter η	0.32	
Opportunity cost of education ξ		.5	
Leisure at school $\bar{l}$		0.40	
Time discount rate θ		0.032	
Interest rate r		0.04	
Wage rate per unit of human capital w		1.47	
Outcomes of the model		**Blacks**	**Whites**
Years of schooling		12.60	13.12
(Observed years of schooling)		(12.74)	(13.50)
Health expenditure (in $1,000)		1.584	1.802
Average lifetime wealth (in $1,000)		100.9	120.1
Average lifetime consumption (in $1,000)		12.500	13.471
Average labor income when working (in $1,000)		18.850	20.059
Average labor supply when working (in hours)		2,272	2,314
Average hourly wage rate (in $)		13.38	13.89

Table 1.2: Sensitivity Analysis

	Parameter values		Outcome of the model	
Parameter description and notation	Baseline	Parameter Intervals	Schooling Years Difference	Medical Expenditure Difference
Time discount rate θ	0.032	[0.030, 0.034]	0.590 (0.038)	$565 ($152)
Leisure at school $\bar{l}$	.40	[0.386, 0.414]	0.610 (0.081)	$589 ($44)
Mortality production parameter μ	1.101	[0.701, 1.501]	0.565 (0.185)	$609 ($300)
Return of education at 9 years of schooling	0.0913	[0.0888, 0.0938]	0.640 (0.034)	$588 ($177)
Opportunity cost of education ξ	0.50	[0.47, 0.53]	0.646 (0.0217)	$589 ($140)
Relative risk averse parameter σ	2.0	[1.74, 2.26]	0.537 (0.203)	$612 ($358)
Mortality production parameter β	0.445	[0.245, 0.645]	0.638 (0.253)	$653 ($451)
Interest rate r	0.04	[0.026, 0.054]	0.890 (0.733)	$624 ($567)

Figure 1.1a: Lifetime Consumption Trajectories

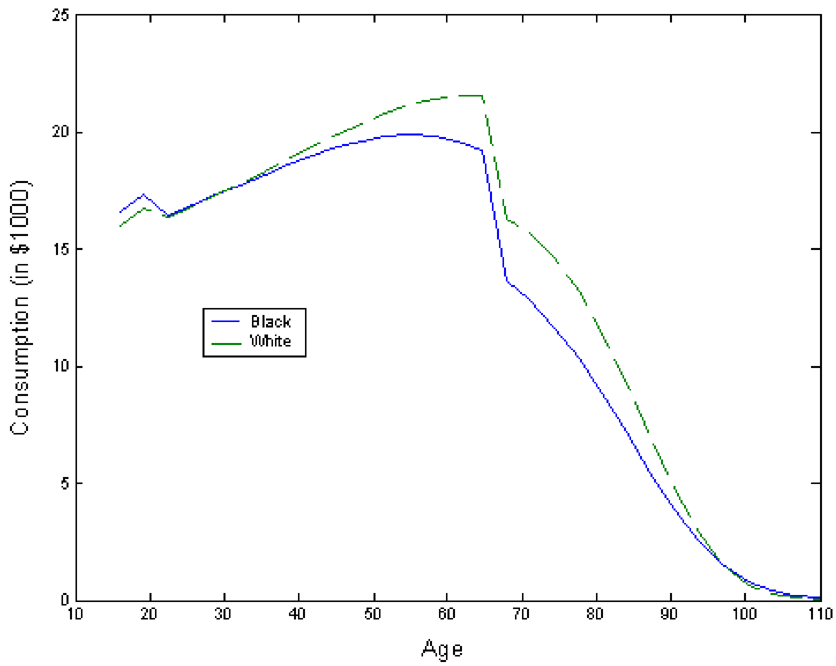


Figure 1.1b: Lifetime Net Income Trajectories

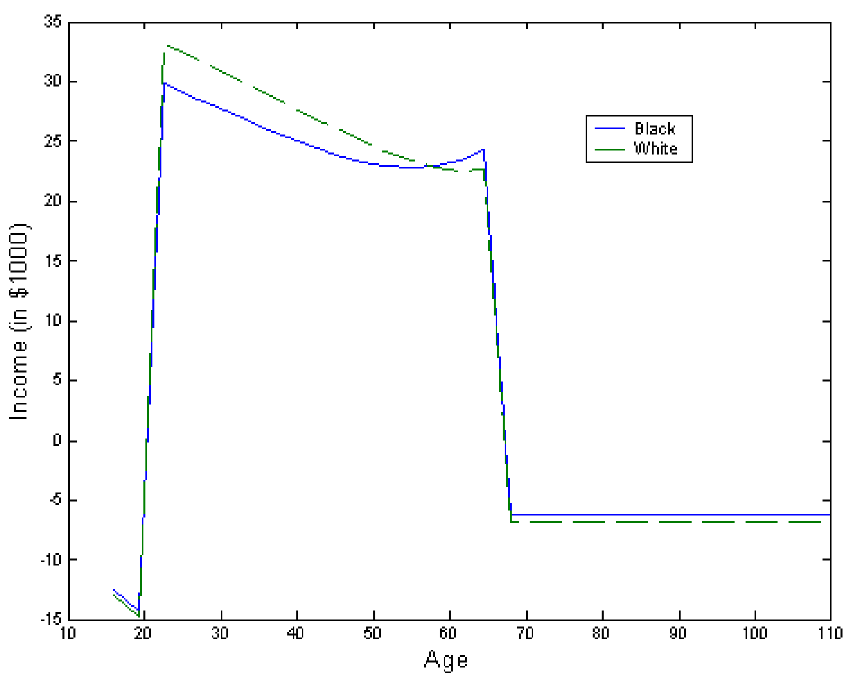


Figure 1.1c: Lifetime Wealth Trajectories

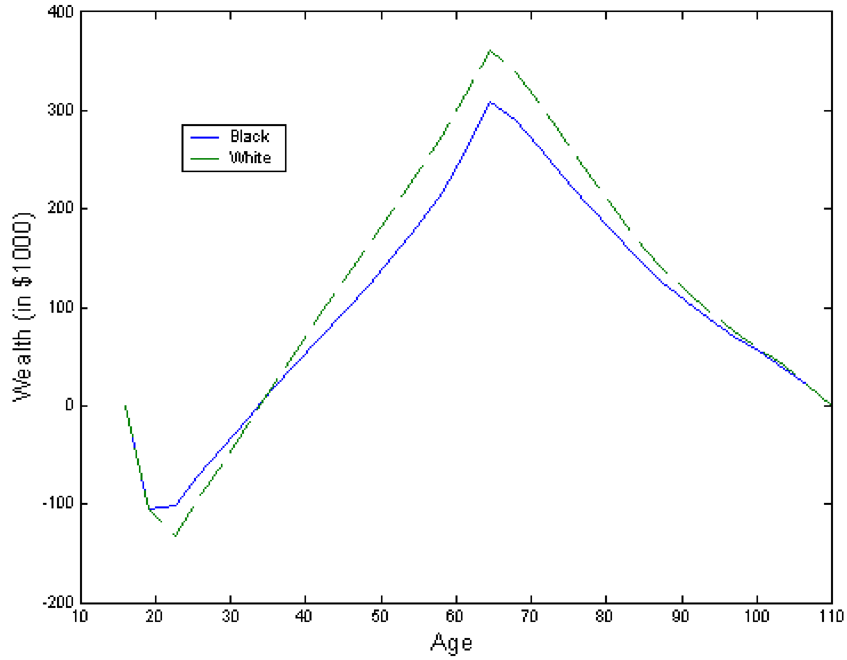


Figure 1.1d: Lifetime Leisure Trajectories

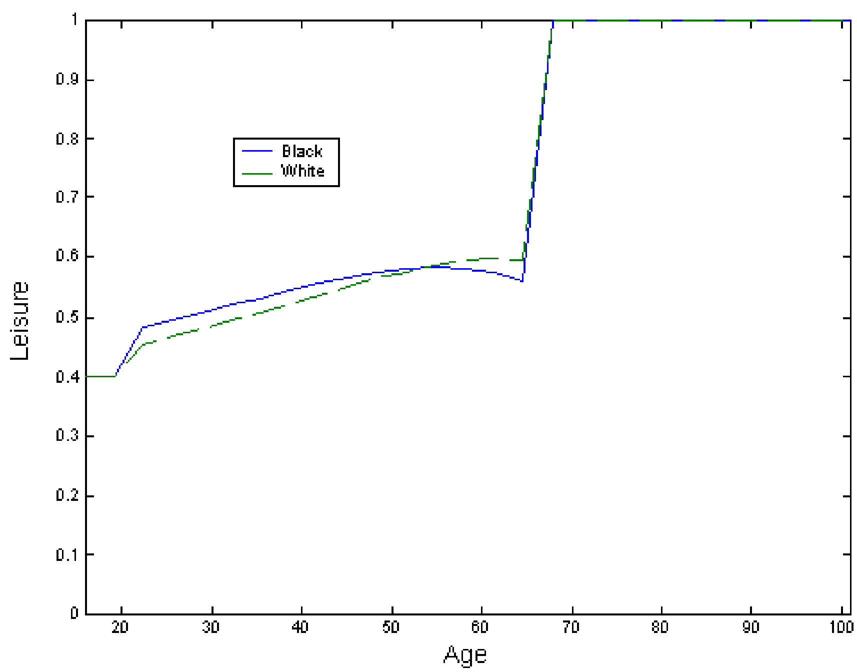


Figure 1.2: Schooling Years as One of the Parameters Varies

Figure 2a: When time discount rate varies

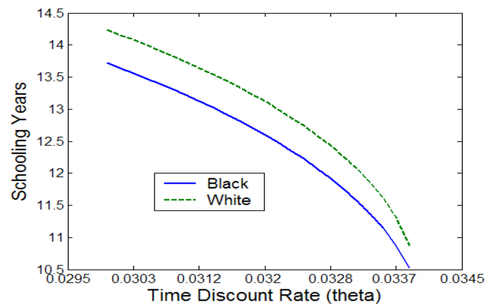


Figure 2b: When leisure in school varies

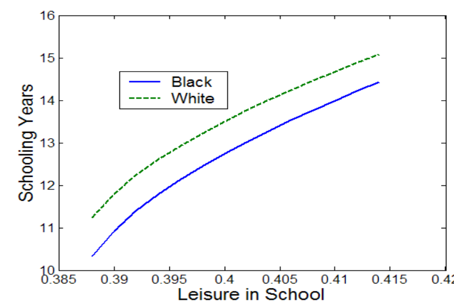


Figure 2c: When mortality production parameter μ varies

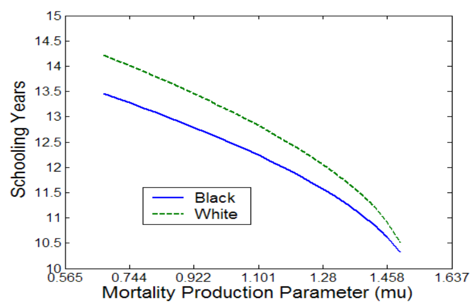


Figure 2d: When return to education varies

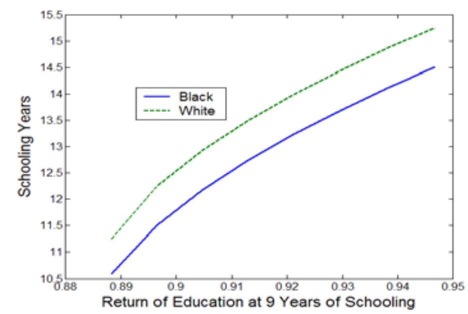


Figure 2e: When cost of education varies

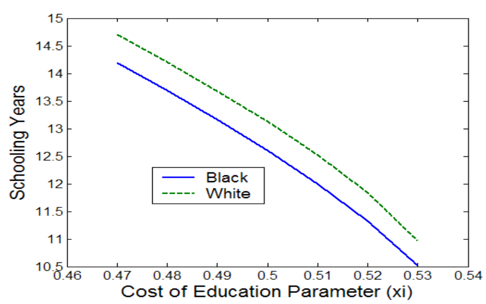


Figure 2f: When risk averse parameter varies

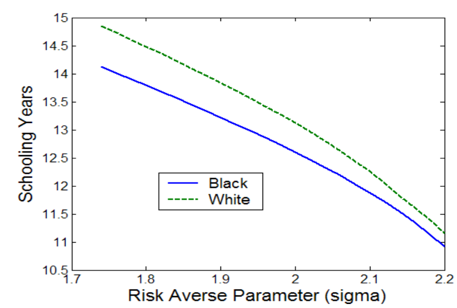


Figure 2g: When mortality production parameter β varies

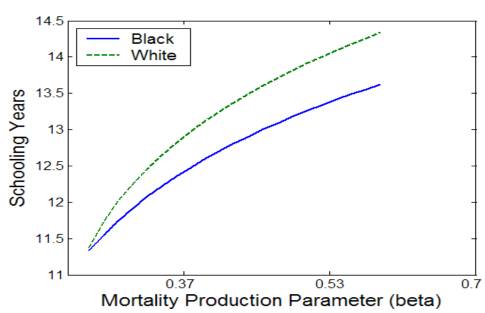
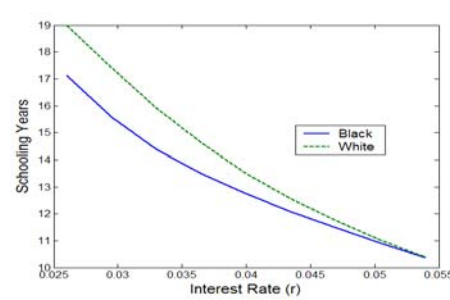


Figure 2h:When interest rate varies



Chapter 2: Estimating Interdependence between Health and Education in a Dynamic Model

2.1 INTRODUCTION

The highly positive correlation between health and education has been well documented in numerous literatures.[1] This finding is robust even after controlling for different measures of socio-economic status, such as income and race, and regardless of whether health levels are measured by mortality rates, self-evaluation of health status, or physiological indicators of health.

The existing literature offers three types of competing theoretical explanations on this correlation. One explanation argues that education increases health by the improvement of economic conditions such as gains in per capita income, and/or by the efficient effect of information, in that more educated people may have more knowledge of health issues (Grossman 1975, Kenkel 1991, Rosenzweig and Schultz 1991). Nevertheless, another controversial explanation argues the reverse causality, i.e. better health results in more education. Healthier students may be more efficient in studying (Perri 1984, Currie and Hyson 1999). Also, better health may increase the demand for education because of the resulting longer life expectancy. Finally, the third explanation argues the existence of a 'third factor,' which affects both health and education in the same direction.[2]

---

[1] See Grossman and Kaestner (1997) for an extensive review.

[2] For example, Fuchs (1982) states that time discount rates could be an explanation for the correlation between health and education: patient people would highly value future income and health and thereafter invest more in education and spend more time and money on activities related to health, while impatient people would invest less in education and health.

Certainly, these three explanations are not mutually exclusive. However, from the public policy perspective, it is important to distinguish between them and to obtain quantitative estimates of their relative magnitudes.

The purpose of this study is to investigate to what extent, in what strength and through which mechanisms health and educational attainment are interrelated. This study provides structural estimates of a dynamic programming model of joint decisions of young men on schooling, work, health expenditure and savings. The structural estimation framework fully imposes the restrictions of the existing theoretical hypotheses on the correlation between health and education, and thus the structural approach provides rigorous interpretations for the parameters that are estimated. Moreover, because the decision rules are explicitly solved by an optimization problem, the estimation permits me to evaluate the impacts of policies related to health improvement, such as financial support in health expenditure, and other monetary incentives to attend college, such as a college tuition subsidy, on an individual's health status, education outcomes and welfare.

A common limitation of previous empirical studies on the correlation between health and education is that they are based on the models of static setting in which education and health are one-shot determined. This method fails to control fully potential biases that may arises as a result of the endogeneity of school enrollment, job participation or consumption choices.[3] For example, a current school attendance decision, which possesses risk as an investment, depends on the probability attached to future work choices and mortality risks. Hence, an individual who has low academic ability and undergoes a bout of sickness, which creates a higher risk of reaping the wage benefits of schooling, may differ systematically from those who have high academic ability and are healthy.

[3] Finding proper and reliable instruments for both health and educational attainment is very difficult in the existing studies.

The static deterministic setting of those models is extended to one in which decision makings are sequential and the environment is uncertain. The model allows for heterogeneity among youth when they reach age 16 in market skills, study skills and health status.[4] These differences may be innate or be a result of prior parental and youth investment behavior or both. The model is estimated using data from the 1979 youth cohort of the National Longitudinal Surveys of Youth (NLSY79). The data provide 16 years of longitudinal information on a representative sample of youth beginning at the age of 16. The model is fit using data for white males on school enrollment, grade transcripts, work status, wages, assets, sickness, and the duration of continuous sickness.

The model contains a number of mechanisms that can account for interactive effects between health and educational attainment. First, health, measured by whether one is sick or not and, if so, the duration of the prior sickness, is assumed to affect the survival rate from now on. Sickness decreases the survival rate and thus decreases the time discount rate, which may consequently result in less school attendance since the individual then may highly value current consumption at the expense of investment in the future.

Second, health is assumed to affect academic performance. The probability for an individual to pass or fail a grade (if he chooses to attend school at all) depends not only on his academic ability but also on his health. Here, health is taken as an important factor affecting the productivity of the study. Third, in a related reason, the model assumes that health affects wages (assuming that job market participation is voluntary) and home productions (assuming that staying home is voluntary), which may account for part of the opportunity costs of school attendance.

---

[4] The sample selection of respondents above age 16 is based on the Fair Labor Standards Act (FLSA). First passed in 1938 and strengthened in a series of succeeding amendments, FLSA severely restricts the use of child labor. According to this legislative regulation, children under the age 16 are strictly restrained to work under the conditions that affect their schooling and health.

Fourth, the model allows the current health status to affect future health status. The individual is assumed to be constantly at risk of sickness. Current health status affects future health because it carries an individual's body and mental information and therefore it will impart into future health. Finally, the model assumes the possibility that education may affect the chance of getting sick, as more educated people are more efficient producers of health.[5]

The estimate of the sickness function indicates that education has a positive effect on the probability of sickness, however the effect is much less significant than the effects of health status and health expenditure. This study also find that health has a substantial effect on an individual's mortality rate, wage, home production, and academic success in school. Indeed, health plays an extremely important role in determining an individual's educational attainment. On average, having been sick before the age of 21 decreases the years of education attained by 1.4.

The estimates of the model are applied to perform two policy experiments: a direct college tuition subsidy and a high school health expenditure subsidy. To assess the efficiency of the policies, These two artificial experiments are allowed to undertake the same amount of per capita cost. The results reveal that a health expenditure subsidy conditional on high school attendance would have a larger impact on educational attainment than a direct college tuition subsidy. In particular, a direct college tuition subsidy will favor healthy individuals, especially those healthy and having low academic ability, while a subsidy of high school health expenditure will favor sick individuals, especially those sick and having high academic ability.

---

[5] The efficiency effect, discussed in detail by Grossman (1999), can take two forms: productive efficiency and allocative efficiency. Productive efficiency pertains to a situation in which more educated obtain a larger health output from given amounts of endogenous (choice) inputs. Allocative efficiency pertains to a situation in which schooling increase information about the true effects of the input on health (Kenkel 2000). Allocative efficiency will improve health to the extent that it leads to the selection of a better input mix.

It should be emphasized that the results of the estimation rest on a strong identification assumption. Most importantly, data limitations require one to make a strong assumption in order to identify the health expenditure that is central to the present model. In fact, the NLSY does not contain direct observations on health expenditure. Rather, the model, in effect, infers the amounts of health expenditures from the individuals' trajectories of asset accumulations, and their choice decisions such as work and school attendance.

A key assumption made to identify the unobserved health expenditure is that only the individuals whose incomes are above a minimum level spend on health. Because a large body of evidence shows that indigent people make very few health expenditures, the minimum level could be interpreted as the income level, below which the individual's primary concern is the consumption of only necessary commodities. Only if the individual's income is beyond the critical level, may he consider spending on health. The minimum income level is exogenous to the individual, although it is estimated as a parameter from the structure model. Therefore, it is possible to identify the health expenditure by comparing the different paths of asset accumulation among high-income groups who spend on health and low-income groups who do not.

The estimation of the model applies the method of generalized indirect inference (GII), recently developed in Keane and Smith (2003). The broadly used methodology at current stage to deal with the dynamic discrete choice model is maximum likelihood (ML) or method of moments (MOM). To estimate the model based on ML or MOM, the econometricians have to evaluate the choice probabilities that the individual could make on each alternative. It is well known that that evaluation of choice probability is computationally burdensome when the number of alternatives is large. This difficulty arises because the choice probability is a high dimensional integral over stochastic factors that affect the utilities that the individual assigns to each alternative. In the present context, the computational problem is formidable for consistent estimation of such models because of severe problems created by unobserved initial conditions and unobserved state variables. Most importantly, the NLSY79 does not contain asset information for 1979-1984 and 1991. Further, the NLSY79 dose not track the transcript record beyond high school.

GII provides a practical simulation-based approach to estimation of dynamic discrete (or discrete/continuous) choice models when there are unobserved variables and many alternatives. This approach builds on the indirect inference. The idea of the indirect inference is to use a descriptive statistical model to summarize the statistical properties of the observed and simulated data from the structural economic model. The method then chooses the structural parameters so that the coefficients of the descriptive statistical model in the simulated data match as closely as possible those in the observed data. Since indirect inference is based on simulated data, it avoids the need to construct the choice probabilities generated by the model. However, the implementation of indirect inference in a discrete choice model encounters a serious problem because of the non-smooth objective function. GII overcomes this obstacle by using a function of the latent utilities as the dependent variable in the descriptive statistical model. As the smoothing parameter goes to zero, this function delivers the discrete choice implied by the latent utilities, thereby guaranteeing consistency.

The choices made for each individual from age 16 to 31 are simulated to implement the GII approach. Individuals differ in their skill endowments, in their health status, and in their schoolings. In each period, the individuals made choices among mutually exclusive and exhaustive alternatives: schooling, work, home, net saving, and health expenditure. The current health statuses and the current incomes associated with work and home have stochastic elements that are known to the individuals at the timing of making decisions but are unknown in the last period. Although the individuals do not know if they will succeed in school before making the decision of whether or not to attend school, they know the probability of passing or failing the grade. Individuals take divergent paths of schoolings, work, home, saving and health expenditures because of the cumulative effects of various shocks and because they have heterogeneous skill endowments and heterogeneous initial health status.

The study is organized as follow. Section 2 presents the model, its basic structure, solution method, estimation method, and parameterizarion. Section 3 describes the data. Section 4 presents the estimation results and describes the policy applications. Section 5 concludes the study.

2.2 MODEL

This section presents the structure of the model with the environment settings as well as the solution and the structural estimation methods. The model corresponds to the decision problem of a young man beginning at age 16. At each period, he decides to enter the labor market, go to school, or to stay at home. In addition, he will decide the amount of health expenditure and saving (or consumption). His current health status is assumed to affect his choice decisions, and in turn that his current choices will affect his future health.

2.2.1 Basic structure

2.2.1.1 Choice set

The element of an individual's choice set at each age t consists of a combination of work participation, school attendance, home, health expenditure, and asset (or saving, and thus consumption).[6] d_t^1 is denoted as the 1×3 choice vector for work, school, and home at period t, with $\sum_{j=1}^{3} d_{j,t}^1 = 1$, where subscript j =1, 2 or 3 corresponds respectively

[6] I discretize saving and health expenditure in order that the individual's choice set is entirely discrete, which increases the tractability of the problem.

to the choice of work, school, and home. More specifically, $d^1_{1,t} = 1$ if the individual chooses to work at period *t*, otherwise $d^1_{1,t} = 0$. Similar are $d^1_{2,t}$ and $d^1_{3,t}$ to $d^1_{1,t}$. Furthermore, the individual will choose among *K* fixed number of discrete levels of saving in excess of interest income, $\{\Delta A^1, \Delta A^2, \ldots, \Delta A^K\}$, where A is the level of asset, and $\Delta A_{t+1} = A_{t+1} - (1+r)A_t$. Denote $1\times K$ vector d^2_t as the decision on the level of saving with $\sum_{k=1}^{K} d^2_{k,t} = 1$, i.e., $d^2_{k,t} = 1$ if ΔA^k is chosen, otherwise $d^2_{k,t} = 0$. Thus, although (excess) savings fall in this range, the feasible asset may grow with age. It is necessary to note that net borrowing is not ruled out in that ΔA may be less than zero. Finally, the individual's choice of health expenditure is divided among the *M* fixed number of discrete levels that are no less than zero: $\{h^1, h^2, \ldots, h^M\}$. Denote $1\times M$ vector d^3_t as the decision on the level of health expenditure with $\sum_{k=1}^{M} d^3_{m,t} = 1$, i.e., $d^3_{m,t} = 1$ if h^m is chosen, otherwise $d^3_{m,t} = 0$. Hence, the number of the individual's choice set at each age *t* is $3\times K\times M$.

2.2.1.2 Environment settings

In order to understand how the individual chooses alternatives corresponding to the current information set and stochastic shocks, it is useful to first describe the environment settings.

Figure 2.1 illustrates the order in which stochastic shocks happen and the timing when the individuals make choices on alternatives. In the beginning of period *t*, an individual's health status (sick or healthy) is known and the random shocks of wage and home production are realized. Then, the individual chooses alternatives among a combination of work, school, home, and the levels of (excess) saving and heath expenditure. Given school attendance, the individual will receive a shock for the grade, which will impact his passing or failing the grade. At the end of period *t*, the agent will get a health shock, which together with his prior educational attainment and current health expenditure will determine his health in the next period *t*+1. At the beginning of period *t*+1, the pattern of the timing and the order that shocks occurred in and alternatives decided at period *t* is repeated.

2.2.1.3 Dynamic programming

At each period *t*, the individual is assumed to maximize the present discounted value of lifetime utility from age 16 (*t* =1) to a known terminal age, *t* = *T*. The value function is given by

$$V_t(\Omega_t) = \mathrm{Max}E\left[\sum_{s=t}^{T} \delta^{s-t} u(c_s) P_{s|t} \middle| \Omega_t \right] \tag{2.1}$$

where E is the expectation operator, Ω_t is the relevant information set known to the individual as he enters decision period *t*, and δ is the subjective time discount factor. $u(c_s) = c_s^{1-\rho} / (1-\rho)$ is the contemporary utility at time *s*. $P_{s|t}$ is the conditional survival rate at period *s* based on the information set at period *t*. The information set includes age, educational attainment (as a proxy for human capital), working experience, health, accumulated asset, and contemporaneous shocks from wage and home production. The maximization of the objective function (2.1) is achieved by choices of the optimal sequence of feasible control variables $\{d_s^l, l = 1, 2, \text{and } 3\}$, given current realizations of the stochastic shocks.

The budget constraint for the individual is given by

$$c_t + \Delta A_{t+1} = w_t d_{1,t}^1 + e_t d_{3,t}^1 - ec * I(edu_t > 12) * d_{2,t}^1 - h_t, \tag{2.2}$$

where w_t is wage, e_t is home production including compensation for nonworking, and h_t is the health expenditure. *edu* is the level of educational attainment and *ec* is the cost of education beyond high school including tuition and the room and board, etc. *I* is the indicator function; it equals one if the argument holds, otherwise it is zero. Attending college or graduate school is assumed to incur monetary outlay. Note that in this study, educational attainment and years of schooling are two different concepts. Years of schooling are the total years that the individual has attended school, while educational attainment is the effective years of schooling, i.e., it is the total years of schooling minus the number of grades that an individual failed.

Health expenditure such as spending on appropriate nutrition, vacation and health clubs affects an individual's survival. The individual's choice decision on health insurance and its subsequent effect on an individual's behavior are not modeled in order to make the model tractable[7]. As stated above, the identification of health expenditure comes from a threshold of income. Only after the income is beyond this critical point will the individual spend on health. More specifically, denote *NIB* as the parameter of income boundary, so therefore, if $rA_t + w_t d_{1,t}^1 + e_t d_{3,t}^1 > NIB$, the health expenditure is positive and is zero otherwise.

Initial conditions include health status in the beginning of the decision horizon, the level of educational attainment and the number of years worked (or working experience) completed by the beginning of the decision horizon, and the level of asset accumulation up to the decision horizon. Both work experience and the level of asset accumulation at age 16 are assumed to be zero.

---

7 Insured and uninsured people show many differences in behaviors related to health, including seatbelt use, diet, and exercise. Moreover, both the supply and demand for insurance depend on health status, which confounds the causal effect between insurance coverage and health. Indeed, credible evidence that access to health insurance causes better health is limited (Newhouse, 1993, Levy and Meltzer, 2001).

2.2.1.4 Probability of sickness

Health status in the next decision horizon is uncertain. The probability for the agent to be sick in the next period, for example, at age *t*+1 depends on his age, his present health expenditure, h_t, his educational attainment, edu_t, and his present health status. Assume that the probability of sickness follows Probit. Define

$$H_{t+1} = \beta_1 age_{t+1} + \beta_2 h_t + \beta_3 edu_{t+1} + D_t\left(\beta_4 + \beta_5 sl_t\right) + \varepsilon_{t+1}^S, \qquad (2.3)$$

where ε_{t+1}^S is serially independent standard normal. The parameter β_3 reflects the idea that more educated people might have efficient knowledge on health issues and thereby refrain from activities that are harmful to health. D_t is an indicator for sickness at age *t*, i.e., $D_t = 1$ if the agent is sick at age *t*, otherwise $D_t = 0$. sl_t is the continuous duration of prior sickness up to age *t*. For example, $sl_t = 0$ if $D_{t-1} = 0$, and $sl_t = 1$ if $D_{t-1} = 1$ and $D_{t-2} = 0$. Then

sick if $H_{t+1} > 0$,

not sick if $H_{t+1} \le 0$.

2.2.1.5 Survival rate

Heath, which is measured by whether the individual is sick or not and if so the duration of the prior sickness, affects the individual's survival rate. Suppose that being sick at age *t* increases an individual's mortality risk, m_t. The mortality rate function, which is based on the life table, is given by

$$m_t \equiv -\frac{P_{t+1} - P_t}{P_t} = \begin{cases} \hat{m}_t e^{\alpha_0 + D_t(\alpha_1 + \alpha_2 sl_t)}, & \text{if} \quad \hat{m}_t e^{\alpha_0 + D_t(\alpha_1 + \alpha_2 sl_t)} < 1 \\ 1, & \text{otherwise} \end{cases}, \qquad (2.4)$$

where α are parameters, $\hat{m}_t$ is the mortality rate of life table at age *t*. Parameter α_2 measures the effect of the duration of prior sickness on the individual's mortality if the individual was sick at the current age. The expression of mortality rate implies that if the agent recovers from a previous period of illness, his current mortality risk will not be affected by his sickness during the previous period. The survival rate at *t*, conditional on being alive at *s*, can thereby be written as

$$P_{s|t} = \begin{cases} \prod_{j=t}^{s-1} \left(1 - m_j\right), & \text{as} \quad s > t \\ 1, & \text{as} \quad s = t. \end{cases} \tag{2.5}$$

It is important to point out that education, income and wealth will affect the survival rate indirectly through their effects on heath expenditure, although they are not included in the mortality function. In fact, NLSY79 does not collect the information on the survival rate in each age. The identification of the survival rate comes from the solution of structural model.

2.2.1.6 Passing or failing a grade

Academic progress is uncertain given school attendance. Assume that whether the individual passes or fail a grade depends on his (unobserved) academic type: high or low intelligence. It is also contingent on his health. The chance to pass or fail the grade follows Probit. Let

$$\Phi_t = \sum_{k=1}^{2} \xi_{0k} I\left(study\,type = k\right) + D_t\left(\xi_1 + \xi_2 sl_t\right) + \varepsilon_t^G \tag{2.6}$$

where ξ are parameters. Summation is over study type, where 1 and 2 indicate high and low skill, respectively. ε_t^G is a serially independent random shock following standard normal distribution.[8] Then

$$\text{pass if } \Phi_t > 0,$$

$$\text{fail if } \Phi_t \leq 0.$$

2.2.1.7 Wage

Assume that wage is a logarithm function of educational attainment (or effective schooling years), edu_t, work skill type (high or low), work experience, ep_t, which is measured by cumulative years worked, age, health and idiosyncratic shock ε_t^w:

$$\begin{aligned}\ln w_t = &\sum_{k=1}^{2} \gamma_{0k} I(skill\,type = k) + \gamma_1 edu_t + \gamma_2 ep_t + \gamma_3 ep_t^2 + \gamma_4 age_t \\ &+ D_t(\gamma_5 + \gamma_6 sl_t) + \gamma_7 I(ep_t = ep_{t-1}) + \varepsilon_t^w\end{aligned} \tag{2.7}$$

where the parameter γ_7 is the adjustment cost if the individual didn't work in the previous period. Unobserved types are incorporated into the wage function in order to reflect the effect of different market skills on wages. Note that the wage rate is standard Mincirian, except for the health terms.

[8] The unobserved random variable may include the individual' level of motivation in study, and the quality of the teacher.

2.2.1.8 Home production

Home production, the reward for remaining home, is unobserved to econometricians, and includes the compensation and any output the individual made. It is assumed that home production depends on an individual's health and is given by

$$e_t = \bar{e} + D_t(\phi_1 + \phi_2 sl_t) + \varepsilon_t^e \tag{2.8}$$

where $\bar{e}$ is constant, and ϕ are parameters. The within-period joint distribution of shocks ε_t^w and ε_t^e is assumed to be serially independent and follows $N(0,\Lambda)$. Since the NLSY79 does not provide the information on home production, to identify the home production I simply assume that only health affect home production.

2.2.2 Solution method

The maximization problem can be set into a dynamic programming problem framework. The value function can be written as the maximum over alternative-specific value functions, each of which obeys the Bellman equation:

$$V_t(\Omega_t;\psi) = \max_{i\in\Gamma}\{V_t^i(\Omega_t;\psi)\} \tag{2.9}$$

where ψ is the vector of parameters of the structural model. Γ is the Cartesian product set of alternatives $Z = d^1 \times d^2 \times d^3$, which consists of $3\times K\times M$ elements. The state space is $\Omega_t = \{edu_t, ep_t, A_t, d_{t-1}^1, d_{t-1}^2, d_{t-1}^3, D_t, sl_t, \varepsilon_t^w, \varepsilon_t^e\}$ (Note that the grade shock, ε_t^G, is not included in the state space, because ε_t^G, as described in the environment settings, is only certain to the individual after the choice decision on school attendance has been made. The health shock, ε_t^S, is known to the individual prior to the choices decisions and its information is reflected in the indication of sickness, D_t.) The alternative-specific value function, $V_t^i(\Omega_t;\psi)$, is given by

$$V_t^i(\Omega_t;\psi) = u^i(\Omega_t;\psi) + \delta(1-m_t)E\left[V_{t+1}(\Omega_{t+1};\psi)\middle|\Omega_t, Z_t^i = 1\right]$$
$$\equiv u^i(\Omega_t;\psi) + \delta(1-m_t)E\max_t\left(\Omega_{t+1};\psi\middle|\Omega_t, Z_t^i = 1\right), \qquad t<T \tag{2.10}$$

$$V_T^i(\Omega_t;\psi) = u^i(\Omega_T;\psi) + \delta(1-m_T)E\left[V_{T+1}^*(\Omega_{T+1};\psi)\middle|\Omega_T, Z_T^i = 1\right], \qquad t=T \tag{2.11}$$

where $u^i(\Omega_t;\psi)$ represents the contemporary utility if the alternative i is chosen (i.e. $Z_t^i = 1$). V_{T+1}^* is the terminal function and will be discussed later. The second equation refers to the notation in Keane and Wolpin (1994) for convenience. More specifically, $E\max_t$ can be written by

$$E\max_t = \Pr_t\left(sick\middle|\Omega_t, Z_t^i = 1\right)E\left[V_{t+1}(\Omega_{t+1};\psi)\middle|\Omega_t, Z_t^i = 1, sick\right]$$
$$+ \Pr_t\left(healthy\middle|\Omega_t, Z_t^i = 1\right)E\left[V_{t+1}(\Omega_{t+1};\psi)\middle|\Omega_t, Z_t^i = 1, healthy\right], \tag{2.12}$$

if school attendance was not chosen, i.e. $d_{2,t}^1 = 0$. And

$$E\max_t = \Pr_t\left(pass\middle|\Omega_t, Z_t^i = 1\right)\Pr_t\left(sick\middle|\Omega_t, Z_t^i = 1\right)E\left[V_{t+1}(\Omega_{t+1};\psi)\middle|\Omega_t, Z_t^i = 1, pass, sick\right]$$
$$+ \Pr_t\left(fail\middle|\Omega_t, Z_t^i = 1\right)\Pr_t\left(sick\middle|\Omega_t, Z_t^i = 1\right)E\left[V_{t+1}(\Omega_{t+1};\psi)\middle|\Omega_t, Z_t^i = 1, fail, sick\right]$$
$$+ \Pr_t\left(pass\middle|\Omega_t, Z_t^i = 1\right)\Pr_t\left(healthy\middle|\Omega_t, Z_t^i = 1\right)E\left[V_{t+1}(\Omega_{t+1};\psi)\middle|\Omega_t, Z_t^i = 1, pass, healthy\right]$$
$$+ \Pr_t\left(fail\middle|\Omega_t, Z_t^i = 1\right)\Pr_t\left(healthy\middle|\Omega_t, Z_t^i = 1\right)E\left[V_{t+1}(\Omega_{t+1};\psi)\middle|\Omega_t, Z_t^i = 1, fail, healthy\right], \tag{2.13}$$

if school attendance was chosen, i.e. $d_{2,t}^1 = 1$.

The elements of the state space evolve according to

$$ep_{t+1} = ep_t + d_{1,t}^1; \tag{2.14.1}$$

$$edu_{t+1} = \begin{cases} edu_t + 1, & \text{if passing the grade} \\ edu_t, & \text{otherwise} \end{cases} \tag{2.14.2}$$

$$A_{t+1} = (1+r)A_t + \sum_{k=1}^{K} \Delta A^k d_{k,t}^2; \tag{2.14.3}$$

$$D_{t+1} = \begin{cases} 1, & \text{if sick} \\ 0, & \text{if not sick} \end{cases} \tag{2.14.4}$$

$$sl_{t+1} = D_t(sl_t + D_t), \tag{2.14.5}$$

and finally, ε_{t+1}'s are serially independent.

Given the finite horizon, the solution method is conducted through backward recursion. The difficulty with this procedure is the well-known "curse of dimensionality" problem. When the dimension of the state space and the choice set are large, the solution of the model becomes formidably difficult in terms of computation time and memory. The high dimensional problem is particularly severe in the present structural model. At each period, the choice set $d^1 \times d^2 \times d^3$ contains 405 (3×15×9) elements.[9] As the decision periods increase, the state space increases exponentially. An approximation method developed in Keane and Wolpin (1994) is adopted to tractably deal with this problem. Specifically, the *E*max approximated function is a polynomial of the state space elements. First, at each period *t*, the *E*max$_t$ function is computed at a randomly selected subset of the state space points by using Monte Carlo integration to simulate the required multivariate integrals. Second, a regression function is estimated as a polynomial in those state space points. Finally, the *E*max values at the non-simulated state space points are interpolated by using the predicted values from the regression.

Finally, the terminal condition has to be specified to solve the maximum problem. The recursion begins at a computationally convenient age, T = 31, to avoid the computational burden of solving the model over a very long horizon. Also, the polynomial form of the *E*max function at that age is used as the terminal condition:

[9] 15 possible values for net asset savings are ± (7500, 5000, 3000, 2000, 1000, 500) and 0, 10000 and 15000. 9 possible values for health expenditure are 0, 250, 500, 750, 1000, 1500, 3,000, 5000, 7,500.

$$
\begin{aligned}
V_{T+1}^{*} = {} & \tau_{01} + \tau_{02} I(\textit{study type is high}) + \tau_{03} I(\textit{work type is high}) + D_{T+1}(\tau_1 + \tau_2 sl_{T+1}) \\
& + \tau_3 edu_{T+1} + \tau_4 edu_{T+1}^2 + \tau_5 A_{T+1} + \tau_6 A_{T+1}^2 + \tau_7 ep_{T+1} + \tau_8 ep_{T+1}^2 \\
& + \tau_9 edu_{T+1} I(\textit{study type is high}) + \tau_{10} edu_{T+1} I(\textit{work type is high}) \\
& + \tau_{11} A_{T+1} I(\textit{study type is high}) + \tau_{12} A_{T+1} I(\textit{work type is high}) \\
& + \tau_{13} ep_{T+1} I(\textit{study type is high}) + \tau_{14} ep_{T+1} I(\textit{work type is high}).
\end{aligned}
\tag{2.15}
$$

The parameters of this terminal function are estimated along with the structural parameters of the model.

2.2.3 Estimation method

The application of GII to estimate the dynamic programming problem (2.9) can be implemented in three steps. The first step is to estimate the descriptive statistical model using the observed data. Denote $\{y_{it}\}_{i=1}^{N}, t = 1,\ldots,T$ as the observed choices and outcomes from the structural model (2.9) with the parameters ψ_0. The observed choices include job participation, school attendance, or remaining at home, while the outcomes include discrete variables of progress in school (passing or failing the grade) and healthy/sick status, and continuous variables of wages and assets. It should be known that because of the missing and unobserved variables, the content of y_{it} may be different over individuals and over periods of time. For example, wages are sometimes unobserved and data for progress in college and graduate school were not collected in the NLSY79, in which case, at some specific period t some individuals' observed data set y_{it} might include wages and/or progress in school, but some might not. In addition, for the years 1979-1984 and 1990, the NLSY79 does not contain asset data, therefore up to age 21,[10] no asset data are in $\{y_{it}\}_{i=1}^{N}, t = 1,\ldots,5$, and after that age some individuals' observed data might include asset information while some might not. The details of the descriptive

[10] The 14-16 years old cohort in 1979 was among 21 –23 years old in 1986 when the data on asset were first collected. Thereby, the youngest age at which some individuals started to have asset information is 21.

statistical model and the selections of dependent and independent variables by ages and by observed data are described in the Appendix A.

Denote the likelihood function associated with the descriptive statistical model as $L(y;z,\Theta)=\prod_{i=1}^{N}\prod_{t=1}^{T} l(y_{it};x_{it},\theta_t)$, where x_{it} is the vector of regressors in the descriptive statistical model for individual *i* at time period *t,* which is described in the Appendix A; y is the set of observed choices, $\{y_{it}\}$; z is the observed exogenous individuals' initial variables, including heath status, educational attainment, working experience (0) and asset (0); Θ is the parameter set $\{\theta_t\}_{t=1}^{T}$. The estimation of the descriptive statistical model using observed data gives the parameters

$$\hat{\Theta}=\arg\max_{\Theta} L(y;z,\Theta) \tag{2.16}$$

The second step is to estimate the descriptive statistical model using the simulated data of choice decisions and outcomes over the decision horizons from the structural model. Given the initial condition *z* and the structural parameters ψ, the structural model can be used to generate *F* statistically independent simulated data sets $\{\tilde{y}_{it}^{f}(\psi)\}_{i=1}^{N}$, where $f=1,\ldots F$, $t=1,\ldots,T$; *N* is the number of observations. The vector of $\tilde{y}_{it}$ corresponds to that of y_{it} in the case that they consist of the same type of elements such as choice decisions (school, work or home, d_t^1), indicators for the passing of a grade and sickness D_t, wages w_t and assets A_t. The generation of $\{\tilde{y}_{it}^{f}(\psi)\}_{i=1}^{N}$ is based on the above described solution method of the simulation and interpolation for computing *E*max. Each of the *F* simulated data sets is constructed using the same set of observed exogenous individuals' initial variable *z*. The difference of each simulated data set results solely from the different sequences of error draws, which are held fixed for different values of the parameter ψ.

Each of the simulated data sets can then be applied to estimate the descriptive statistical model. However, it is not practical computationally to simply plug in the simulated discrete variables into the descriptive statistical model because of the non-smooth objective function (actually its surface is a step function).[11] Applying the idea of GII proposed in Keane and Smith (2003), a series of latent utility is used to substitute the discrete choice variables and related variables. More specifically, the function is used

$$\tilde{\tilde{d}}_{1,t}^{1}(\psi;\lambda)=\frac{\sum_{j\in\Xi_1}\exp\left(V_t^j(\Omega_t;\psi)/\lambda\right)}{\sum_{j\in\Gamma}\exp\left(V_t^j(\Omega_t;\psi)/\lambda\right)} \tag{2.17}$$

in place of simulated $\tilde{d}_{1,t}^{1}$, where Ξ_1 is a subset of alternative set Γ and consists of all the alternatives in which job participation is chose. Because the latent utilities are smooth functions with parameters ψ, $\tilde{\tilde{d}}_{1,t}^{1}(\psi;\lambda)$ is a smooth function of ψ. Moreover, as the smooth parameter λ goes to 0, $\tilde{\tilde{d}}_{1,t}^{1}(\psi;\lambda)$ goes to 1 if an alternative with job participation has the highest latent utility and to 0 otherwise.

Similarly, the function

$$\tilde{\tilde{d}}_{2,t}^{1}(\psi;\lambda)=\frac{\sum_{j\in\Xi_2}\exp\left(V_1^j(\Omega_t;\psi)/\lambda\right)}{\sum_{\Gamma}\exp\left(V_1^j(\Omega_t;\psi)/\lambda\right)} \tag{2.18}$$

is used in place of simulated $\tilde{d}_{2,t}^{1}$, where subset Ξ_2 consists of all the alternatives in which school attendance is chosen. As the smooth parameter λ goes to 0, $\tilde{\tilde{d}}_{2,t}^{1}(\psi;\lambda)$

[11] The reason for the difficulty in practice is discussed in detail in Keane and Smith (2003): 'small changes in the structural parameters ψ will cause the simulated data jump discretely and such a discrete change caused the parameters of the descriptive model fit to the simulated data to jump discretely. This jump, in turn causes the metric of distance between the descriptive models estimated on the observed and simulated data to jump discretely too. The algorithms to deal with the minimization of a non-smooth function perform very poorly.'

goes to 1 if an alternative with school attendance has the highest latent utility and to 0 otherwise.

Wages are observed if and only if the individuals worked during that period. To make the simulated wage match the observed wage, the observed wage is applied for those individuals who worked at that period, and set the wage to zero for those individuals who did not work at that period. $\tilde{\tilde{d}}_{1,t}^{1}(\psi;\lambda)\tilde{w}_{it}(\psi)$ is used in place of the simulated wage $\tilde{w}_{it}(\psi)$. Since both $\tilde{\tilde{d}}_{1,t}^{1}(\psi;\lambda)$ and $\tilde{w}_{it}(\psi)$ are smooth functions of ψ, the estimated parameters of the descriptive statistical model using the simulated data are also smooth functions of ψ. Moreover, as the smoothing parameter λ goes to 0, $\tilde{\tilde{d}}_{1,t}^{1}(\psi;\lambda)\tilde{w}_{it}(\psi)$ goes to $\tilde{w}_{it}(\psi)$ if an alternative with job participation choice has the highest latent utility and to 0 otherwise.

Furthermore, because the indicator of sickness is a discrete variable, it need be substituted by a continuous function. Function

$$\tilde{\tilde{D}}_{t+1}(\psi;\lambda)=\frac{\exp\left(H_{t+1}(\psi)/\lambda\right)}{1+\exp\left(H_{t+1}(\psi)/\lambda\right)} \tag{2.19}$$

is used in place of simulated $\tilde{D}_{t+1}$. Thus, as the smooth parameter λ goes to 0, $\tilde{D}_{t+1}(\psi,\lambda)$ goes to 1 if H_{t+1} > 0 and to 0 otherwise.

Finally, according to the same reason for the discrete variable of sickness, the continuous function $\exp\left(\Phi_t(\psi)/\lambda\right)/\left[1+\exp\left(\Phi_t(\psi)/\lambda\right)\right]$ is used in place of the indicator for passing a grade.

Denote $\left\{\tilde{\tilde{y}}_{it}^{f}(\psi;\lambda)\right\}_{i=1}^{N}$, t = 1, …, T, and f = 1, …, F, as the modified simulated data smoothed by using the functions of the latent utilities. The descriptive statistical model then can be estimated using each of the simulated smoothed data to obtain the following parameters

$$\tilde{\tilde{\Theta}}_f(\psi;\lambda) = \arg\max_{\Theta} L\left(\tilde{\tilde{y}}^f(\psi;z); x, \Theta\right). \tag{2.20}$$

Let the average of the estimated parameters be $\tilde{\tilde{\Theta}}(\psi;\lambda) = \sum_{f=1}^{F} \tilde{\tilde{\Theta}}_f(\psi;\lambda)/F$. As the sample size N goes to large and the smooth parameter λ goes to small (zero), $\tilde{\tilde{\Theta}}(\psi;\lambda)$ converges to a nonstochastic "binding" function $\mathrm{H}(\psi)$ (Gourieroux, Monfort, and Renault (1993) and Keane and Smith (2003)). The next step of the GII is to get an estimate $\hat{\psi}$ of the structural parameters so as to make $\mathrm{H}(\hat{\psi})$ and $\hat{\Theta}$ as close as possible.

The third step is to estimate the structural parameter ψ by minimizing a metric function, which measures the distance between $\hat{\Theta}$ and $\tilde{\tilde{\Theta}}(\psi;\lambda)$. In the present context, The likelihood ratio is adopted as the metric function, which is used in Keane and Smith (2003). In particular,

$$\hat{\psi} = \arg\max_{\psi} L\left(y; z, \tilde{\tilde{\Theta}}(\psi;\lambda)\right) \tag{2.21}$$

In regards to the practical algorithm for the estimation of the structural model, this study adopts the two-step approach following Keane and Smith (2003). The details of choosing the number of simulated data sets F and the smoothing parameter λ of each step are described in the Appendix A.

Finally, the generation of an individual's skill endowments is described as follows. The endowment skills at age 16 are assumed to be unobserved to the econometricians. But the population proportions of the skill types are known.[12] Denote

[12] Keane and Wolpin (1994, 1997) undertake the same assumption. Some literatures use the Armed Forces Qualifying Test (AFQT) as a measure of IQ or endowment skill (Neal and Johnson, 1996; Cameron and Heckman, 1998, 1999). The reason that this analysis does not adopt AFQT is based on the following basics: (1) AFQT reflects not only the innate endowment but also the parents and own investments in skills up to the time of the test. But, due to the age distribution of the samples in the NLSY79, small portion of the individuals took the test prior to age 16. (2) Given that each individual is characterized by two skills

the type portions of high ability for studying and high skill for working as ro_1 and ro_2, respectively. An individual's skill types can be simulated by random draws from the uniform distribution between 0 and 1. For example, if an individual *i*'s drawn number corresponding to the study type is less than ro_1, the individual is labeled as having high academic ability. Otherwise the individual is labeled as having academic ability. At each simulated data *f*, the individual's skill types are generated independently from the random draws.

2.3 DATA

The dataset used for the structural model estimation is from the 1979 youth cohort of the National Longitudinal Surveys of Youth (NLSY79). The NLSY79 contains extensive information about the individuals' employment, education, health, income and asset. An original 12,686 individuals, a nationally representative sample, were interviewed each year from 1979 to 1994. After 1994, the interviews switched to every other year.

The analysis is based on the sample of the white males who were age 16 or younger as of October 1, 1977 and never sworn into active military service. Basically, this study rules out the choice of military. This study follows each individual in the sample from the first year they reach age 16 as of October 1 that year through September 30, 1993.

(studying and working), one-dimensional AFQT obviously could not adequately represent a two-dimensional skill.

2.3.1 Health

In each survey year, the NLSY79 asked the individuals a standard set of health questions. The focus of these questions was on the health problems that affected the respondent's ability to work.[13] In each year, if the respondents were not currently working, they were asked if their health would (was) prevent(ing) them from working, and the rests of respondents working currently were asked if their health limited the type and the amount of work they could do. If a health limitation was reported, the NLSY79 then probed for the month and year the health limitation began.

This study uses the answers to these questions to construct the health variables for this analysis.[14] An individual was classified as sick in a given year if a health limitation was reported at that year. The construction of the sick duration variable is based on the information of the individual's reported date of when the sickness began. The difficulty in constructing health variables is that a large portion, around thirty percent, of the self-reported sick duration in the NLSY79 did not match the preceding self-reported sickness. For example, some respondents reported that the sickness began at some earlier point, say, two years ago, but no reported health limitation could be found during the last 2 years. This could be because that no surveys were conducted for these respondents at that periods, or because that the respondents had not been aware of the sickness until the health limitations developed into a serious problem affecting their lives. To solve this inconsistent problem, I check the subsequent self-reported answers to health questions, also searching for references to the specific ailments. If the respondents kept reporting the same health problems and the same date the heath limitation began, I then use this

---

[13] People currently in school may intend to report a healthy situation and therefore cause bias. But it is not clear what kind of bias it will make. To deal with this issue is beyond this study scope.

[14] More specific details on the health ailment were asked in the NLSY79 if the individuals gave an affirmative answers to being limited in either the kind or amount of work they could do because of health. However, the afterwards tremendous computation of the structural model restraints me from considering these individual's specific health problems.

information to update the prior sickness variables. If the specific health problem only reported once but the duration was longer than one year during the entire time of the survey, I simply classify the respondent as sick only during that reported year.

In the constructed health data, twenty one percent of the respondents report at least one illness during the 16 years of surveys. The average duration is 2.28 years. Figure 2 shows the percentage of respondents who reported sickness at each age from 16 to 29.[15] At the early age of 16, 4.14% of the white males in the samples consider themselves sick. Over the subsequent 15 years, the percentage of the respondents reporting an illness increases steadily, peaking at the age of 29 with 5.17%.

2.3.2 Schooling, work, or home

At each interview date, the NLSY79 asked the respondents about their enrollment status, the highest grade attended and completed, school-leaving dates, and the dates that diplomas and degrees were received. An individual is considered as attending school during the year if the individual reported enrollment in school at the time of the survey and did not report dropping out of school during that year in the subsequent surveys.

Employment data in the NLSY79 include the beginning and ending dates of all jobs, hours worked on each job (to the calendar week), and salary paid on each job. An individual is considered to have worked during the year if the individual reported working at least 1000 hours, i.e. at least 20 hours per week on average for 50 weeks.

An individual is considered to be at home during the year if the individual neither was enrolled in school nor worked during the year. Note that some individuals would be

---

[15] The figure ends at age 29 instead of 31. The percentages of sickness report at age 30 and age 31 are 4.81% and 5.18%, respectively. A dip at age 30 and the breaking of the increase trend may come from the shrinking of sample size. During the annual survey from 1979 to 1993, 98.4% of the original respondents reached their age 28, however, only 73% and 43.6% of the respondents reach their 30 and 31 years old, respectively.

classified as being at home if they worked during the year but did not work at least 1000 hours.

Table 2.1 presents the choice distributions of white males by age for the whole sample and for the sickness data. The sickness data are cumulative, i.e., at each age *t*, it consists of the individuals who have reported being sickness at least once up to and including age *t*. The initial number of individuals is 1,062 at age 16. From age 16 to age 29, the number declines slightly as a result of sample attrition such as natural decease. The fall of the sample size during age 30 and 31 is because some samples never reach their older ages during the survey periods. Overall, there are 15,972 person-periods in the whole sample dataset and 2,198 person-periods in the sickness dataset. As the table shows, the choice decisions on school attendance, job participation, or remaining home are highly dependent on the individual's health. Compared to the individuals in the whole sample, individuals in the sickness data at each age have a smaller percentage of school attendance and in contrast they have a larger percentage of remaining home. Moreover, although a slightly larger percentage of individuals in the sickness data worked during the first three ages, a relatively smaller percentage of sick individuals worked after that. More specifically, 11.56% of the individuals among the sickness data attended school, 42.81% worked, and 45.63% remained at home. The corresponding percentages for the individuals among the whole sample are 25.34%, 54.46%, and 20.20%. Furthermore, the relative difference in the percentage of school attendance between the two data sets increases during the normal schooling ages. For example, at age 16, the percentage of individuals attending school while having been sick is 81.82% (i.e., 93.6% of the average 87.38%), but at the normal high school graduation age of 18, that percentage drops to 38.55% (i.e., 77.2% of the average 49.95%), and at the normal college graduation age of 22, it drops to 11.51% (i.e., 60.01% of the average 19.18%). Additionally, the propensity to work increases monotonically over the first 11 years of both data sets, followed by slight fluctuations over the last 5 years.

Tables 2.2 and 2.3, which respectively show one-period transition rates for the whole sample and sickness data in the process of making choices, reveal substantial state dependence for health status. The row percentages describe the transition percentages conditioned on the prior state, and the column percentages show transition percentages conditioned on the succeeding state. State persistence is revealed in the Tables 2.2 and 2.3. A large majority of the individuals who enrolled in school in the last period will enroll currently, however over 73% of the whole sample and less than 60% of sickness sample will make such a decision. Also, individuals who worked or remained home last period have the same pattern as those who enrolled, and thus will work or be home at this period. Moreover, the individuals who have been sick before are more likely to remain home if they were at home during the last period, and are less likely to work if they worked last period.

2.3.3 Passing or failing grades

The NLSY79 collected the information of the high school transcripts during 1980, 1981, and 1983 for those respondents who were 17 years of age or older including those who dropped out from high school, and who were expected to complete high school in the United States. The transcript data gathered for each of up to 64 courses include grade level at which the course was taken, a code for the high school course, a grade for each course based on a 0 to 4.0 scale, and the credits received.

To conduct the dummy variable for an individual passing or failing a grade, I simply assume that an individual failed a grade if and only if the individual failed over a half of the courses taken in that grade. This assumption implies that each course is equally important for assessing the progress in school. Because no clear guideline or uniform standard exists for weighting the grade of each course, the current assumption provides a feasible way to evaluate the performance of study in school. Further, a course is classified as failure for an individual if its grade is lower or equal to F.

Table 2.4 shows the percentages of failing in high school by grade for the whole sample and for the sickness data. The sickness data consists of the respondents who reported health limitation at least once up to and including the year the grade was taken. As shown in the Table 2.4, health limitation significantly influences the outcome of study. For the same reason that health problems could limit the type and amount of work an individual could do, health also affects the grades an individual could attain by affecting the efficiency of an individual's effort at study. Specifically, the possibility of failing a grade declines as the grade level becomes higher, from 13.9% in grade 9 to 3.63% in grade 12. The declining trend in grade failures may reflect the fact that some students could not complete high school and dropped out of school before graduation because of bad grades, health problems, or both. In addition, the possibility of failing a grade for the individuals who had been sick is more than twice than the average, except for the grade 9 in which the failing possibility is about 1.5 times higher.

2.3.4 Wage and asset

The real wages used in this analysis are based on 1984 price level. On average, an individual who was sick during the surveys of 16 years will make his wage decrease about 10%. More specifically, the average wage in the whole data is \$20,752 with a standard deviation of \$47,535, in contrast to an average wage of \$18,731 in the sickness data, with a standard deviation of \$11,367.

Beginning in 1985, the NLSY79 launched a much larger wealth section. Up to 20 questions about a variety of asset and debt holdings were asked to the respondents at each subsequent interview, except for 1991.[16] The asset items used in this analysis include (i)

[16] The wealth questions were eliminated in 1991 round of survey because of the budgetary restrictions.

residential property, (ii) cash savings, stock and bond portfolio, etc., (iii) real estate, assets in the business, and farm operation, (iv) automobile, (v) mortgage debt, property debt, and other accumulated debt, (vi) other assets each worth individually more than \$500, (vii) other debts over \$500. Together these variables provide an analog of the net worth of the assets of each respondent. In actuality, the NLSY79 did not distinguish the assets of the respondents from those of their spouses. The respondents were asked the amount of the asset owned or owed by the respondents and/or the spouses. To construct the net asset belonging to the respondent in a survey year, I check the respondent's answer to the question: "are you currently married and is your spouse listed on the household enumeration?" If the answer is no, his net asset is the value calculated from the above items; otherwise, his net asset is cut in half.

Table 2.5 and Table 2.6 show the asset distribution by age for the whole sample and the sickness data, respectively. The earliest age with reported assets is 21 because the asset data were not collected until 1985, and the age range selected in this analysis is from 14 to 16 in 1979. Given the small size of the observations and possible measurement errors, outlier asset levels, which are dependent on age, are deleted.[17] As shown in the tables, both mean and median net assets in sickness data are smaller than those in the whole data, reflecting the substantial and persistent influence of sickness on the accumulation of assets. The prevalent dependence of assets on health is also verified by the proportions of the negative net assets, which are higher in the sickness data from the ages 22 to 31. In addition, tables 2.5 and 2.6 indicate that assets increase with age. Between the ages of 21 and 31 in the whole sample and sickness data, the mean net assets increase by 4.13 and 2.90 times, respectively, while the median net assets increase by 3.75 and 3.63 times, respectively. Moreover, the median net assets are, on average, less

[17] In total, 107 extremely large and small net asset observations are gotten rid of from the whole data, while 34 from the sickness data.

than half of the mean levels. This reflects the positively skewed nature of the asset distribution.

2.4 ESTIMATION RESULTS

2.4.1 Parameter estimates

The parameter estimates are reported in table 2.7. The standard deviations are in parentheses. In total, the number of parameters is 50. The parameters are estimated to fit the sequential choices of 15,972 person-period observations, out of which 2,198 had been sick at least once through the 16-year surveys. The choice set at each period consists of decisions on school attendance, job participation, or staying at home, decisions on net asset savings and health expenditure, and received wages.

The estimated parameters for the mortality rate function show that a healthy individual's mortality is 1.5% lower from the value of life table, while the mortality rate for an individual who has experienced sickness with zero duration is 21 times larger than the life-table mortality and the mortality rate rises to 22 times as large as the life-table mortality if the sickness duration is 3.5 years. As for the survival rate, being sick at age 16 with 0 duration decreases the survival rate between age 16 and 30 by 2% from 98.4% to 96.4%.

The estimates for passing or failing a grade indicate that health and academic skill endowment have a significant effect on the success of an individual in each grade. Among the individuals with high academic ability who account for 86.5% of the population, the probability for a healthy individual to pass a grade is 97.6%, while the corresponding probability for a sick individual with zero duration and a sick individual with 3-year duration drops to 91.2% and 90.88%, respectively. On the contrary, of the individuals with low academic ability, the probability of passing a grade is 75.4% if he is healthy and 53.3% if he is sick with zero duration. In general, health plays a more important role for an individual of low academic ability in whether or not he will pass the grade than for an individual of high academic skill. Specifically, the passing probability of an individual with high academic skills will decrease by 6.4% as a consequence of a sickness, whereas the probability of failure with low academic skill will decrease by 22.8%.

The estimates for wage equation parameters reveal that sickness reduces wages by 16%. In addition, individuals with high working skill (approximately 59% of the population) earn about 30% more than low working type individuals if other characteristics are the same. Furthermore, the job adjusting cost and the returns for education, experience and age match the empirical estimate results: the absence of work in the last period decreases wages by 13%; an additional year of education increases wages by 10%; an additional year of experience increases wages by 11% in the first year and after that an additional year of experience increases wages by $10.85\text{-}2*0.4*ep_t$ percent, which implies that wages reach a peak after 13.6 years (other things kept constant); wages decrease 0.7% per year.

With respect to the home production function, estimated parameters show that sickness reduces the home production by \$2716, and an additional year of duration reduces the home production \$368. The average home production for a healthy individual is \$9689. Additionally, wage shock and home production shock are negatively correlated with the correlation coefficient -0.3816.

Tables 2.8 and 2.9, based on the estimated parameters, report the probabilities of being sick at selected ages, health expenditures, health statuses, and educational attainments. As the tables show, both health expenditure and heath status have significant effects on the possibility of sickness, however the effect of education on the possibility of sickness is relatively much smaller. Table 2.8 takes a fixed 10-grade educational attainment. It shows that a healthy individual has around a 50 percent chance of getting sick if he does not spend on any health activity, while a sick individual has a high probability of more than 87 percent of getting sick. At age 16, a \$500 health expenditure helps a healthy individual reduce the probability of sickness 81%, from 48% to 9%, and the relative reducing strengths for sick individuals fall between 49% and 31%, decreasing with the increase of sickness duration. At more advanced ages, the effect of health expenditure drops slightly, for example, aged 30, with a \$500 health expenditure, an individual's probability of sickness is reduced 78% if healthy and 44% if sick. Therefore, relative to health expenditure and health status, the aging effect is pretty small for the possibility of sickness.

The estimated effect of education on the probability of sickness in table 2.9 is divided into grades 8, 12 and 16, which represent the education level of pre-high school, high school graduate, and 4-year college graduate, respectively. Table 2.9 shows that education has a negative effect on the probability of sickness, although the effect is much less significant than heath expenditure and health status. Specifically, if there is no health expenditure, the probability of sickness for a 20-year-old and healthy individual at eighth grade is 51%, and at grades 12 and 16, the corresponding probabilities of sickness drop to 48% and 45%, respectively.

Finally, the coefficient of relative risk aversion is 0.8043 and preference discount factor is 0.9795, which are consistent with values in the literature. The estimated cost of education beyond high school is \$4, 328 per year. And, the net income boundary is -\$585, below which the health expenditure is zero.

2.4.2 Within-sample fit

With the estimated parameters, the validation of the model can be tested by the within-sample fit. Based on a simulation of 8,000 individuals, table 10 compares the predicted and actual values of selected state variables divided by the sickness and the whole data. As can be seen, the model accurately matches the mean level of completed schooling years in the whole data. However, the disaggregation of completed schooling years shows that the model overstates the proportion of those having completed twelve years of schooling (high school) and understates the proportion of those having completed sixteen years of schooling (college). In the sickness data, the model predicts a slightly higher mean level of completed schooling years and overstates the proportion of those having completed twelve years of schooling.

The model fits the mean percentage of decision choices on work, school and home quite well, except that it overstates the mean fraction of school attendance in the sickness data and understates the mean fraction of work decision in the whole sample data. A further fit comparison on the predicted and actual age patterns of school attendance, work decision and home decision for the whole sample and sickness data is illustrated in figures 2.3 and 2.4.

In fitting the percentage of those failing a grade, the model predicts a relatively stable percentage of failure in each of the grade level: around 5% in whole set of data and 16-17% in the sickness data. Moreover, the predicted probability of failure in each grade is over 11% higher in the sickness data than in the whole data.

With respect to the asset fit, the model captures the broad increasing pattern of age. Figure 2.4 displays the predicted and actual mean asset by age. It is clear that the predicted mean levels of assets are quite closer to the actual levels for the whole sample data than for the sickness data.

As predicted by the model, the mean health expenditure in the sickness data is 5.4% larger than in the whole set of sample data. This is because that sick individuals have to spend more on health to reduce the chance of being sick in proceeding periods, and healthy individuals can spend small amounts on health activity and still maintain a relatively low probability of sickness. Figure 2.5 shows the predicted and actual percentages of sick individuals respectively among the simulated date and the whole sample data, which are around 4.1% to 5.2%. The largest and smallest gaps between predicted and actual sick percentages are 0.29% at age 30 and 0.02% at age 27, respectively. Moreover, the age pattern of health expenditure and the percentage of zero health expenditure are portrayed in figure 6. It is shown that the mean health expenditure increases by age, from $783 at age 16 to $952 at age 31, an average increase of 1.34% per year. Concurrently, the ratio of zero health expenditure increases from 0 in the first four years (i.e. ages 16 to 19) to 0.94% at age 30. Note that according to the model assumption, as the individual's net income is lower than the boundary, -$549, his health expenditure is zero. The increase trend of the percentage of zero health expenditure implies the dispersion of assets and earnings.

2.4.3 Initial health status and education effects

As has been observed, an individual's initial characteristics have a significant effect on his future behaviors of alternative choices, which will subsequently determine his health, educational attainment and wealth. It is interesting to investigate how the education, health and welfare are related to initial levels of completed education and health status at the age of 16.

Table 2.11 reports the simulation results of initial health status effects on selected variables, conditional upon initial schooling. Approximately 5% of individuals completed 10 years or more schooling by age 16 in the observation sample. As seen in the table, initial health status is an important determinant of education, survival probability, asset, health expenditure and lifetime welfare. Moreover, the effects of initial health limitations are more substantial for individuals with lower levels of education than for individuals with higher level of education. For instance, illness at age 16, on average, decreases the average level of education at age 30 by 0.35 year for individuals with initial schooling of nine years or less, whereas by 0.27 year for individuals with initial schooling of ten years or more. Moreover, the decrease in the probability of survival at age 30, resulted from the illness at age 16, is 2.8% for those with low initial education compared to 1% for those with higher levels of education. Finally, due to the health limitation at age 16, the mean present value of lifetime utility decreases 13% for the individuals with low initial education and 11% for those with high initial education, respectively.

Table 2.11 indicates that initial education has a significant effect on the selected variables, and the magnitude of the initial education effect seems larger than the initial health status effect. But we cannot conclude consequently that the education effect is more important than the health effect in determining educational attainment and welfare before we find a unified measure to quantify their strengths.

2.4.4 Policy application

In this section, I evaluate the efficiency of policies aimed at decreasing inequality. The first policy experiment is a direct college tuition subsidy and the second policy experiment is to subsidize health expenditure during high school. The two experiments will incur the same amount of per capita cost. Therefore, by comparing the outcomes of the artificial experiments I am able to answer the question which policy is more efficient: spending on education or spending on health?

2.4.4.1 College tuition subsidy

Table 2.12 reports the distribution effect of a $2,100 college tuition subsidy, which is about 50% of college tuition. Although the subsidy is limited to college students, it will also affect the individuals' decisions before entering college because they will anticipate it before making their decision to enter college. The people are divided into two groups: those having been sick at least once before age 21 (12.2% of the population before the subsidy) and those who have remained healthy before age 21 (87.8% of the population before the subsidy). Also, people are classified by their endowment types based on the estimated parameters of population type ratio: high ability in both school and work (type 1), high ability in school and low ability in work (type 2), low ability in school and high ability in work (type 3), low ability in both school and work (type 4). For convenience, the baseline results without subsidy are listed.

As expected, college tuition subsidy increases the levels of state variables including educational attainment, length of years spend in college, assets, and present value of lifetime utility. The average highest schooling years completed increases by 0.42 years from 13.39 to 13.81 years. And the mean years in college increase by 0.35 years from 1.85 to 2.20 years. The mean value of assets at age 30 increase 18% from $19,134 to $22,608. The mean expected present value of lifetime utility at age 16 increases 10.7% from 185.6 to 197.4. Finally, the percentage of those having faced significant sickness at age 20 decreases 0.9% from 12.2% to 11.3%.

As seen, the college tuition subsidy has a smaller effect on a sick person than on a healthy person. In particular, educational attainment changes little for type 3's and type 4's who have experienced at least one bout of sickness by the age 20. Specifically, the private gain of welfare from the subsidy is smaller for the sick person than for the healthy person. The mean present value of lifetime utility increases 5.6% for those experiencing sickness, compared to 11% for healthy people.

In this case, the per capita cost of college tuition subsidy is $2,247, if shared by all of the individuals. However, the gains are very different across types and heath status, up to age 20. Overall, type 1's and type 2's experience greater gains from the program, who have a significant large college attendance regardless of the subsidy. In addition, healthy people gain more with respect to sick people.

2.4.4.2 High school health expenditure subsidy

Table 2.13 explores the effect of a \$778 health expenditure subsidy for high school students, which can only be used as the health expenditure. With the \$778 health expenditure subsidy, the per capita cost of program is \$2,247, which is the same amount as the per capita college tuition subsidy.

As shown, the average highest year of schooling completed increases by 0.53, which is 0.11 year more than with the college tuition subsidy. The mean years spent in college increases by 0.44, a little larger than with the college tuition subsidy. In addition, the mean assets at age 30 are almost the same as in the case of the college tuition subsidy. The overall welfare has a tiny increase with respective to the college tuition subsidy program.

The gain distribution is much different in this case. Gains of sick and low endowment people improve substantially. This could be explained by the two reasons. First, health limitation decreases the possibility of passing a grade, and graduating from high school is the only path to attending college. Hence, a college tuition subsidy is not as attractive to those who anticipate a small probability of passing a grade. However, high school heath expenditure provides a direct channel for this population to gain from the subsidy. Second, for those people who would go to college even without the tuition subsidy, the benefits are most because of the level effect of the subsidy. But, for those who are induced to attend college, the benefits incurred from the marginal effect, i.e. marginal indifference between college attendance and other options.

2.5 CONCLUSION

In this study I structurally estimate a dynamic model of schooling, work, health expenditure and saving choices over the life cycle using 16 years of data from the NLSY79. The structural estimation framework fully imposes the restrictions of the existing theoretical hypotheses on the correlation between health and education. The model's estimates imply that an individual's initial health status has a substantial influence on an individual's educational attainment and expected survival probability. Indeed, health plays an extremely important role in determining an individual's educational attainment. On average, having been sick before the age of 21 decreases the level of educational attainment by 1.4 years. Policy experiments based on the model's estimates indicate that a health expenditure subsidy conditional on high school attendance would have a larger impact on the educational attainment than a direct college tuition subsidy. In particular, a direct college tuition subsidy will favor healthy individuals, especially those healthy and having low academic ability, while a subsidy of high school health expenditure will favor sick individuals, especially those sick and having high academic ability.

Figure 2.1: Stochastic Shocks and Decisions

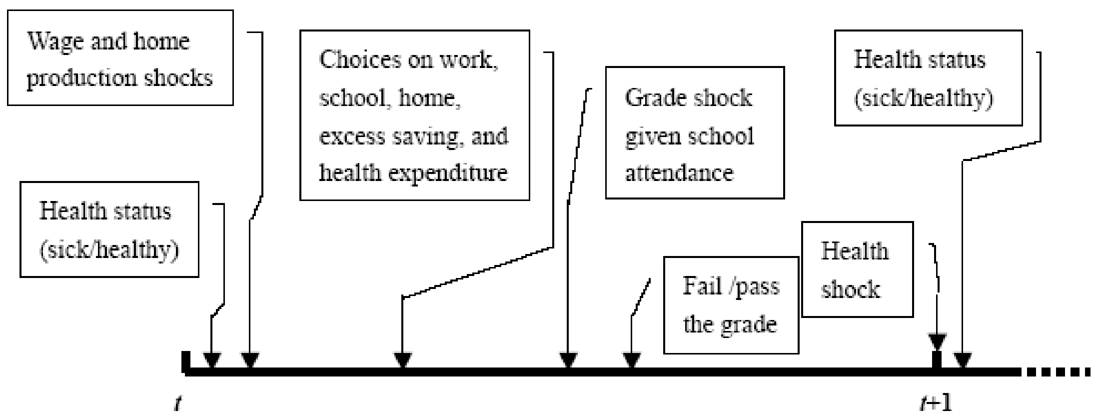


Figure 2.2: White Male Sickness Reports

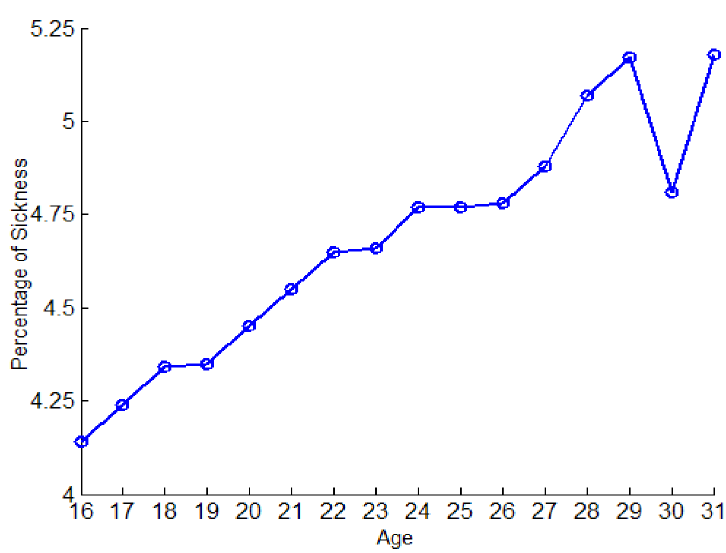


Figure 2.3a: Predicted and Actual Mean Percent Choice Selections by Age: Whole Data

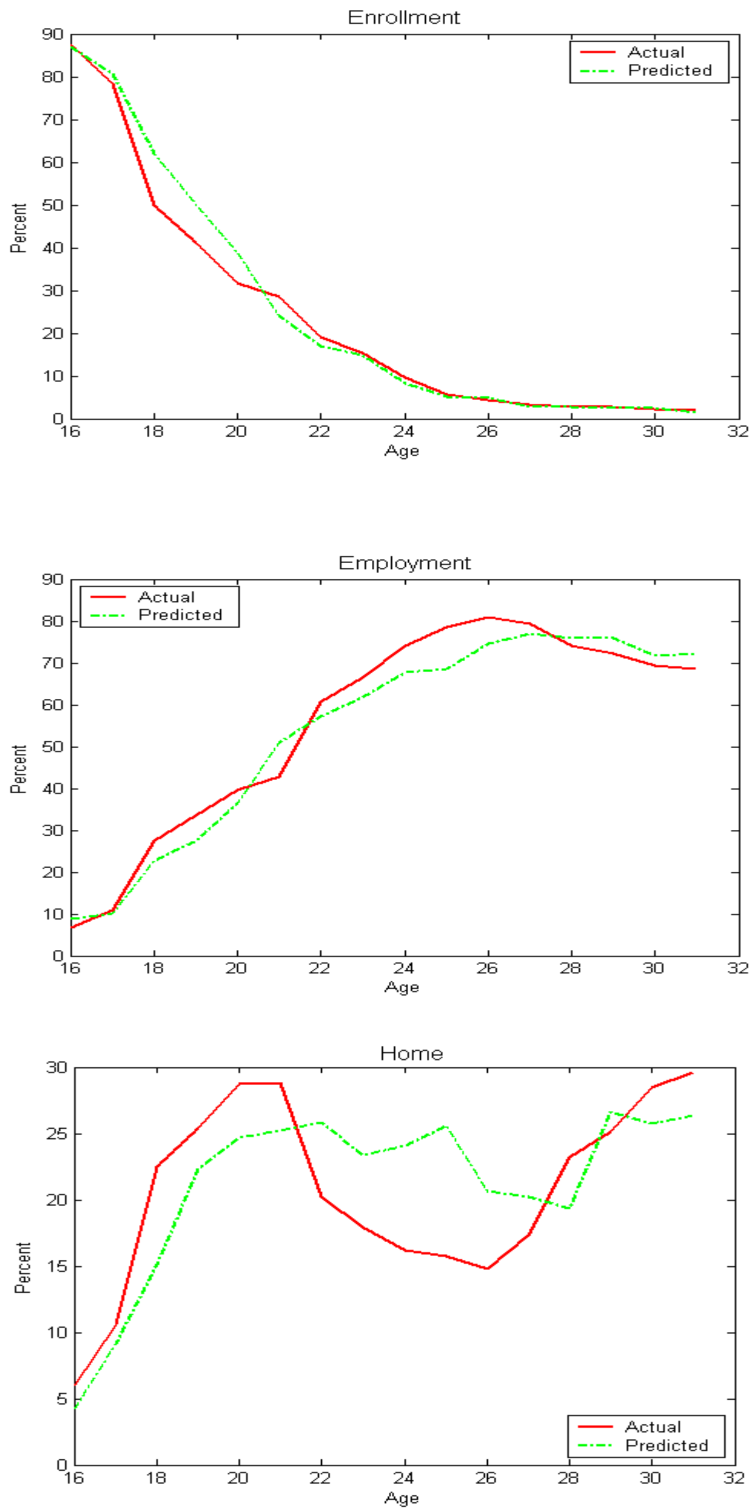


Figure 2.3b: Predicted and Actual Mean Percent Choice Selections by Age: Sickness Data

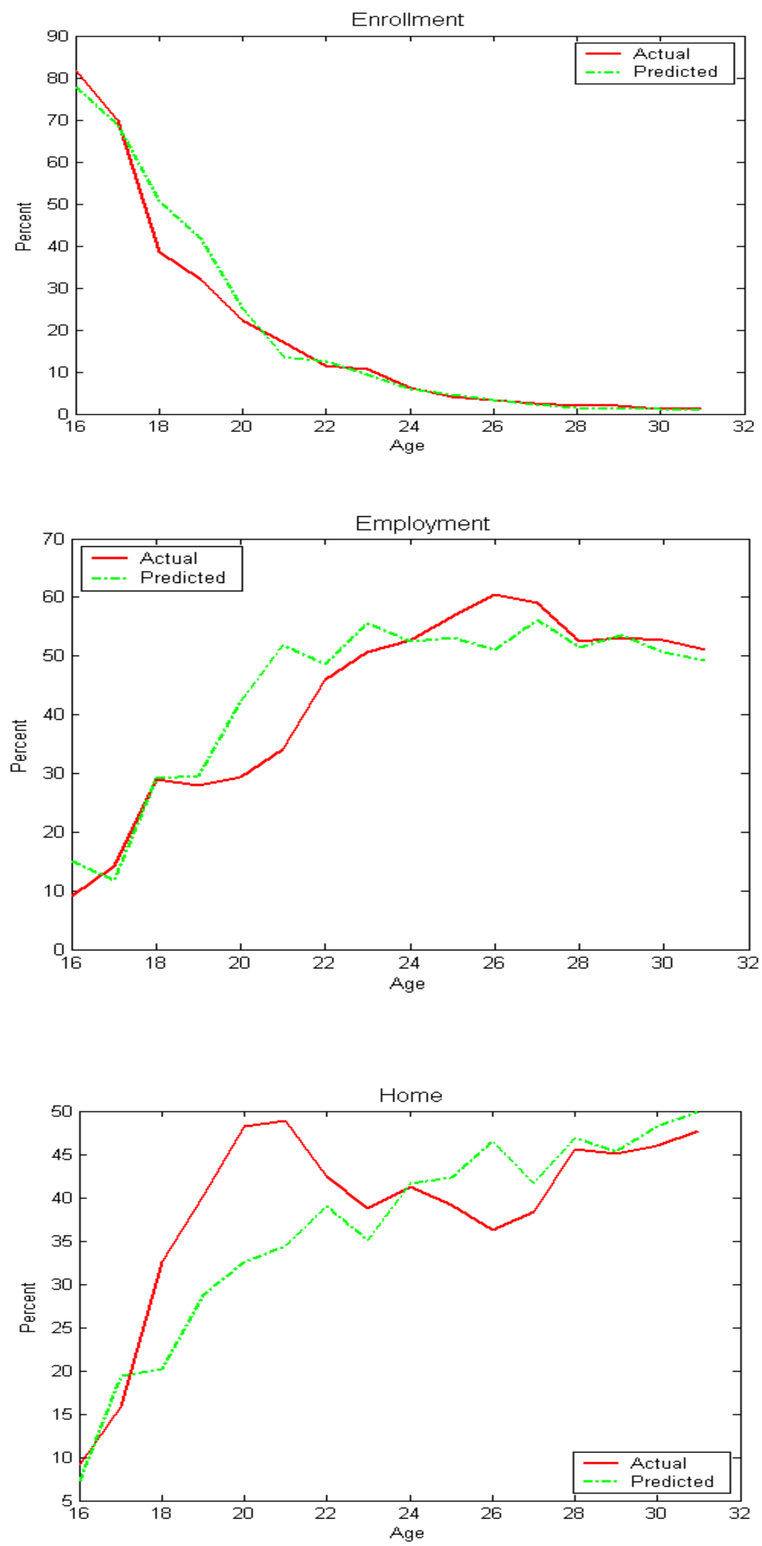


Figure 2.4a: Predicted and Actual Mean Assets by Age: Whole Data

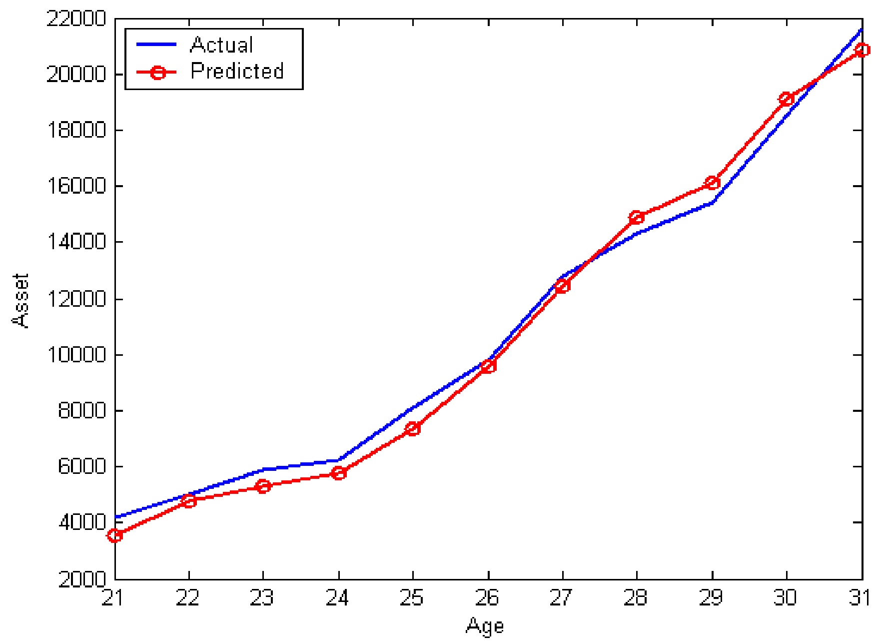


Figure 2.4b: Predicted and Actual Mean Assets by Age: Sickness Data

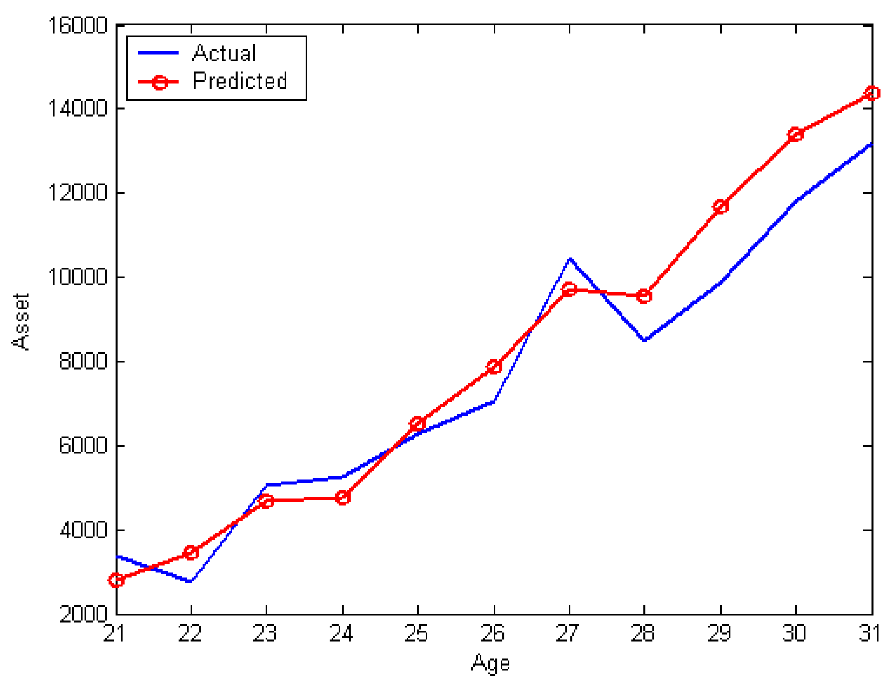


Figure 2.5: Predicted and Actual Sick Percentage

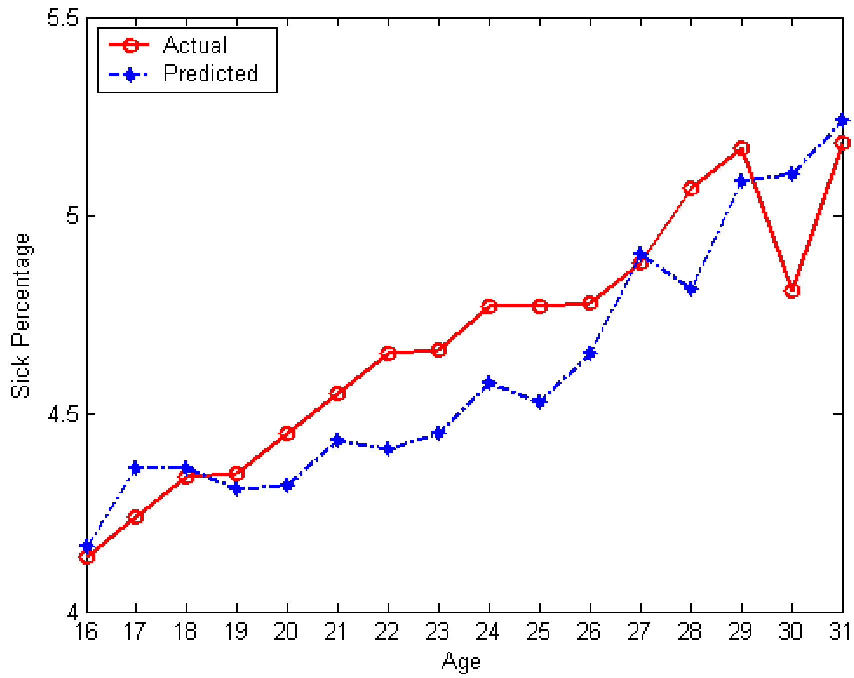


Figure 2.6: Predicted Health Expenditure and Percentage of Zero Health Expenditure

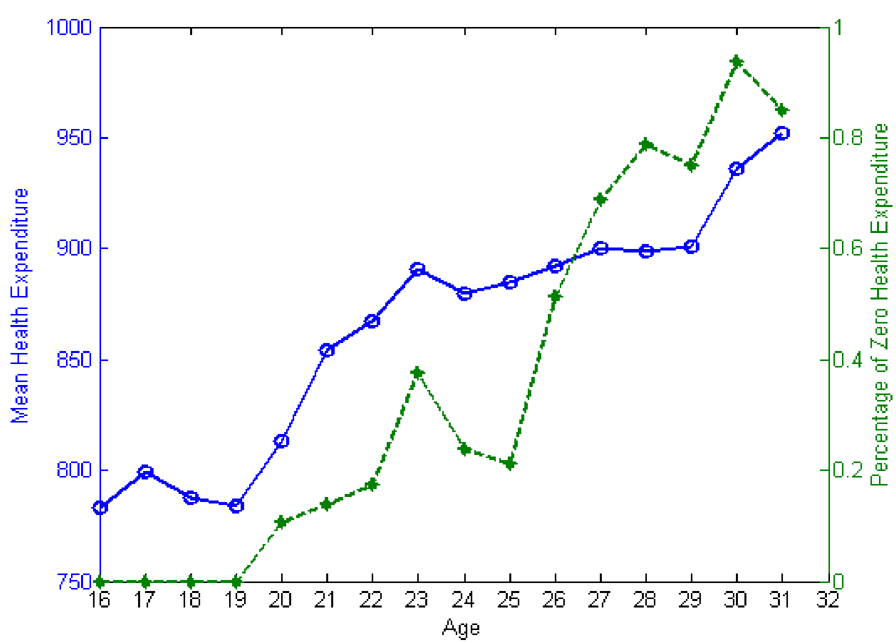


Table 2.1 Choice Distribution: White Males Aged 16-31

	Sickness Data				Whole Data			
Age	School	Work	Home	TOTAL	School	Work	Home	TOTAL
16	**81.82**	**9.09**	**9.09**	**100**	**87.38**	**6.59**	**6.03**	**100**
	(36)	(4)	(4)	(44)	(928)	(70)	(64)	(1062)
17	**69.87**	**14.29**	**15.87**	**100**	**78.42**	**11.02**	**10.56**	**100**
	(44)	(9)	(10)	(63)	(832)	(117)	(112)	(1061)
18	**38.55**	**28.92**	**32.53**	**100**	**49.95**	**27.57**	**22.47**	**100**
	(32)	(24)	(27)	(83)	(529)	(292)	(238)	(1059)
19	**31.96**	**27.84**	**40.21**	**100**	**41.12**	**33.46**	**25.43**	**100**
	(31)	(27)	(39)	(97)	(435)	(354)	(269)	(1058)
20	**22.32**	**29.46**	**48.22**	**100**	**31.63**	**39.68**	**28.69**	**100**
	(25)	(33)	(54)	(112)	(334)	(419)	(303)	(1056)
21	**17.05**	**34.11**	**48.84**	**100**	**28.53**	**42.65**	**28.82**	**100**
	(22)	(44)	(63)	(129)	(301)	(450)	(304)	(1055)
22	**11.51**	**46.04**	**42.44**	**100**	**19.18**	**60.59**	**20.23**	**100**
	(16)	(64)	(59)	(139)	(202)	(638)	(213)	(1053)
23	**10.53**	**50.66**	**38.81**	**100**	**15.49**	**66.54**	**17.97**	**100**
	(16)	(77)	(59)	(152)	(163)	(700)	(189)	(1052)
24	**6.06**	**52.73**	**41.21**	**100**	**9.64**	**74.14**	**16.22**	**100**
	(10)	(87)	(68)	(165)	(101)	(777)	(170)	(1048)
25	**3.98**	**56.82**	**39.20**	**100**	**5.63**	**78.63**	**15.74**	**100**
	(7)	(100)	(69)	(176)	(59)	(824)	(165)	(1048)
26	**3.30**	**60.44**	**36.26**	**100**	**4.3**	**80.88**	**14.82**	**100**
	(6)	(110)	(66)	(182)	(45)	(846)	(155)	(1046)
27	**2.63**	**58.95**	**38.42**	**100**	**3.25**	**79.43**	**17.32**	**100**
	(5)	(112)	(73)	(190)	(34)	(830)	(181)	(1045)
28	**1.96**	**52.45**	**45.59**	**100**	**2.87**	**73.97**	**23.16**	**100**
	(4)	(107)	(93)	(204)	(30)	(773)	(242)	(1045)
29	**1.86**	**53.02**	**45.12**	**100**	**2.68**	**72.15**	**25.17**	**100**
	(4)	(114)	(97)	(215)	(28)	(754)	(263)	(1045)
30	**1.24**	**52.80**	**45.96**	**100**	**2.19**	**69.33**	**28.48**	**100**
	(2)	(85)	(74)	(161)	(17)	(538)	(221)	(776)
31	**1.16**	**51.16**	**47.68**	**100**	**1.94**	**68.47**	**29.59**	**100**
	(1)	(44)	(41)	(86)	(9)	(317)	(137)	(463)
TOTAL	**11.56**	**42.81**	**45.63**	**100**	**25.34**	**54.46**	**20.20**	**100**
	(254)	(941)	(1,003)	(2,198)	(4,047)	(8,699)	(3,226)	(15,972)

Note: Percentages and number of observations.

Sickness data at age *t* consist of individuals who have been sick at least once up to and including age *t*.

Table 2.2: Transition Matrix: Whole Data

White Males Aged 16-31

	Choice (t)		
Choice (t - 1)	School	Work	Home
School:			
Row %	73.49	12.98	13.52
Column %	92.59	9.25	17.61
Work:			
Row %	2.31	86.98	10.7
Column %	3.71	78.97	17.77
Home:			
Row %	4.26	23.92	71.82
Column %	3.7	11.78	64.62

* Observation No.: 14910

Table 2.3: Transition Matrix: Sickness Data

White Males Aged 16-31

	Choice (t)		
Choice (t - 1)	School	Work	Home
School:			
Row %	58.7	15.38	25.97
Column %	83.09	6.25	10.66
Work			
Row %	1.81	82.3	15.89
Column %	5.95	77.9	16.15
Home			
Row %	3.49	19.25	77.26
Column %	11.22	16.72	72.05

* Observation No.: 2154

Table 2.4: Percentage Failing for Grades 9, 10, 11 and 12

White Males Aged 16-31

Grade	**9**	**10**	**11**	**12**
Whole data	13.9	6.08	5.54	3.63
	(374)	(954)	(903)	(799)
Sickness data	20.69	20.93	19.57	7.61
	(35)	(72)	(71)	(67)

* Number of observations with transcript report are in parentheses

Table 2.5: Asset Distribution: Whole Data

White Male Aged 21 - 31

Age	Median	Mean	Std	Max	Min	No. Obs.	Percent Negative
21	1,931	4,209	6,404	55,330	-15,296	230	9.8
22	2,248	5,019	8,262	80,524	-14,753	497	11.2
23	2,752	5,883	10,581	115,630	-12,703	921	16.4
24	2,863	6,263	12,507	176,972	-31,618	911	16.7
25	3,590	8,082	16,071	196,907	-36,624	907	15.3
26	4,003	9,833	20,235	209,874	-43,152	938	16.6
27	5,237	12,803	22,458	227,072	-43,722	677	16.7
28	5,565	14,294	26,456	247,706	-33,388	607	15.0
29	7,443	15,424	27,621	262,705	-37,028	438	12.9
30	8,628	18,501	35,369	298,728	-21,211	589	11.6
31	9,168	21,599	48,360	338,994	-24,756	351	10.7

Table 2.6: Asset Distribution: Sickness Date

White Male Aged 21 - 31

Age	Median	Mean	Std	Max	Min	No. Obs.	Percent Negative
21	1,333	3,389	6,306	16,927	-8,035	29	6.7
22	2,058	2,737	4,091	19,434	-7,402	67	20.9
23	2,566	5,064	8,042	36,585	-8,714	130	17.7
24	2,654	5,257	10,030	61,999	-13,719	141	17.0
25	3,000	6,289	9,488	52,133	-10,518	148	17.6
26	3,545	7,054	12,002	62,358	-11,312	160	20.0
27	4,886	10,452	17,390	93,206	-6,415	114	18.2
28	3,481	8,470	14,398	69,612	-12,197	103	18.5
29	3,703	9,898	15,695	84,883	-12,583	97	17.5
30	5,036	11,823	18,375	77,389	-9,347	118	16.1
31	6,169	13,203	23,483	96,098	-8,479	65	12.3

Note: 1984 dollars.

Sickness data at age *t* consists of individuals who reported health limitation

Table 2.7: Estimates of the Model

Mortality Rate Function:	
constant α_0:	-0.0143 (0.00026)
sickness α_1:	3.0566 (0.8965)
interaction between sickness and duration α_2:	0.0136 (0.0032)
Pass/Fail the Grade:	
high study type ξ_{01}:	1.9743 (0.1690)
low study type ξ_{02}:	0.6940 (0.9629)
sickness ξ_1:	-0.6245 (0.2533)
interaction between sickness and duration ξ_2:	-0.0047 (0.0226)
Wage Function:	
high working type γ_{01}:	1.4967 (0.0649)
low working type γ_{02}:	1.2043 (0.579)
educational attainment γ_1:	0.1027 (0.0034)
experience γ_2:	0.1085 (0.0094)
experience squared/100 γ_3:	-0.4008 (0.0789)
age γ_4:	-0.0069 (0.0036)
sickness γ_5:	-0.1624 (0.0473)
interaction between sickness and duration γ_6:	-0.0032 (0.0045)
no working at last period γ_7:	-0.1324 (0.0201)
Home Production Function:	
constant $\bar{e}$:	9689.1 (6245.3)
sickness ϕ_1:	-2715.8 (1376.5)
interaction between sickness and duration ϕ_2:	-368.4 (148.36)
Sick Probability:	
age β_1:	0.0085 (0.0038)
health expenditure β_2:	-2.5694 (0.0489)
educational attainment β_3:	-0.0187 (0.0069)
sickness β_4:	1.2041 (0.3568)
interaction between sickness and duration β_5:	0.1060 (0.0058)
Type Ratio:	
high ability in study ro_1	0.8605 (0.1547)
high skill in work ro_2	0.5859 (0.257)

Estimates of the Model (Cont.)

Terminal Value Function:	
constant τ_{01}:	6.0259 (2.6101)
high study type τ_{02}:	0.1002 (0.0258)
high work type τ_{03}:	0.1011 (0.0326)
sickness τ_1:	-0.547 (0.247)
interaction between sickness and duration τ_2:	-0.0582 (0.0265)
educational attainment τ_3:	5.409 (2.068)
educational attainment squared /100 τ_4:	2.3054 (0.216)
asset τ_5:	0.1594 (0.0231)
asset squared / 1e5 τ_6:	-0.000181 (0.0025)
experience τ_7:	1.1541 (0.269)
experience squared /100 τ_8:	0.1182 (0.146)
interaction between education and high study type τ_9:	0.1664 (0.589)
interaction between education and high work type τ_{10}:	0.1001 (0.0698)
interaction between asset and high study type τ_{11}:	0.001018 (0.0263)
interaction between asset and high work type τ_{12}:	0.000602 (0.0025)
interaction between experience and high study type τ_{13}:	0.001031 (0.006)
interaction between experience and high work type τ_{14}:	0.003028 (0.0024)
Error	
standard deviation of wage σ_w:	0.5137 (0.0698)
standard deviation of home production σ_e:	8.1867 (3.694)
correlation σ_{we}:	-1.6049 (0.1895)
Preference Discount Factor δ:	0.9795 (0.2793)
Coefficient of Relative Risk Aversion σ:	0.8043 (0.3691)
Education Cost *ec*:	\$4328(1569.2)
Net Income Boundary *NIB*:	-\$584.8 (178.25)

Note: standard errors are in parentheses.

Table 2.8: Estimated Sick Probabilities in Percentage by Age, Health expenditure and Health Status

			Sick			
Age	Health Expenditure	Healthy	0-year duration	1-year duration	3-year duration	5-year duration
16	0	47.97	87.56	89.60	92.94	95.38
	$250	24.40	69.52	73.13	79.64	85.1
	$500	9.08	44.77	48.98	57.39	65.48
	$750	2.4	21.95	25.21	32.42	40.36
	$1,000	0.44	7.83	9.50	13.6	18.77
	$1,750	0	0.04	0.06	0.12	0.25
25	0	51.02	89.06	90.92	93.91	96.08
	$250	26.87	72.15	75.59	81.73	86.81
	$500	10.40	47.80	52.03	60.37	68.26
	$750	2.86	24.28	27.71	35.22	43.35
	$1,000	0.55	9.02	10.86	15.34	20.9
	$1,750	0	0.05	0.08	0.16	0.31
30	0	52.71	89.83	91.59	94.41	96.42
	$250	28.29	73.56	76.91	82.84	87.69
	$500	11.19	49.50	53.72	62	69.76
	$750	3.15	25.62	29.15	36.81	45.03
	$1,000	0.62	9.73	11.68	16.37	22.15
	$1,750	0	0.06	0.09	0.18	0.35

* Eucation attainement is 10 grades
** in 1984 dollar.

Table 2.9: Estimated Education Effect on Sick Probability

Grade	Health Expenditure	Healthy	Sick: 0-year duration	Sick: 3-year duration
8	0	50.81	88.96	93.85
	$500	10.31	47.60	60.17
	$1,000	0.54	8.93	15.22
12	0	47.83	87.49	92.89
	$500	9.03	44.63	57.26
	$1,000	0.43	7.78	13.53
16	0	44.86	85.88	91.82
	$500	7.87	41.69	54.31
	$1,000	0.35	6.75	11.97

* Age is 20.

** in 1984 dollar.

Table 2.10: Predicted and Actual State Variables

	Sickness Data		Whole Data	
	Predicted	Actual	Predicted	Actual
Years of schooling:				
Mean highest schooling years completed	12.63	12.58	13.39	13.40
Percent 12 schooling years completed	83.29	75.68	87.08	83.71
Percent 16 schooling years completed	13.52	13.06	17.71	25.80
Mean percentage of employment	41.86	42.81	48.81	54.46
Mean percentage of school attendance	20.54	11.56	28.65	25.34
Mean percentage of home	37.60	45.63	22.54	20.20
Percent grade fail *:				
Grade 9	17.32	20.69	5.61	13.90
Grade 10	17.66	20.93	5.85	6.08
Grade 11	16.44	19.57	5.22	5.54
Grade 12	16.17	7.61	4.97	3.63
Mean assets at age **:				
21	2805	3389	3548	4209
24	4767	5257	5796	6263
27	9722	10452	12466	12803
30	13404	11823	19134	18501
Mean health expenditure***	896.6	--	850.5	--

Note: Based on 8,000 simulated individuals.

The sickness data include all the individuals who have been sick during the 16-year periods.

* In this case, the sickness data consists of cumulative individuals who reported sickness by the specified grade.

** The age-t sickness data consists of cumulative individuals who have been sick up to age t.

*** The health expenditure under the sickness data is specified to the sick period.

Table 2.11: Initial Health Status Effects by Initial Schooling

	Healthy at Age 16	Sick at Age 16
	Initial Schooling Nine Years or Less	
Mean education attainment at age 30	13.17	12.82
Mean percent survival probability at age 30	97.36	94.60
Mean asset at age 30	17,676	14,822
Mean health expenditure by age 30	834	892
Expected present value of lifetime utility at age 16	186.4	162.5
	Initial Schooling Ten Years or More	
Mean education attainment at age 30	14.60	14.33
Percent survival probability at age 30	98.12	97.08
Asset at age 30	28,654	24,586
Mean health expenditure by age 30	1030	1072
Expected present value of lifetime utility at age 16	224.6	204.1

Note: Based on a simulation of 8,000 persons.

Table 2.12: Effect of a $2100 College Tuition Subsidy on Selected State Variables*

Characteristics	All	Sick up to and including Age 20**					Healthy up to and including Age 20				
		All Types	Type 1	Type 2	Type 3	Type 4	All Types	Type 1	Type 2	Type 3	Type 4
Average highest schooling years completed:											
No subsidy	13.39	12.17	12.35	12.72	10.12	10.24	13.56	13.64	14.03	12.0	12.26
Subsidy	13.81	12.55	12.76	13.16	10.16	10.32	13.97	13.92	14.45	12.81	13.15
Mean years in college:											
No subsidy	1.85	0.73	0.65	1.12	0.05	0.07	2.01	2.19	2.34	0.43	0.61
Subsidy	2.20	1.1	1.15	1.44	0.05	0.07	2.34	2.26	2.82	1.32	1.58
Asset at age 30											
No subsidy	19,134	11,725	12,464	13,237	5,474	4,794	20,163	21,047	21,369	14,099	13,597
Subsidy	22,608	13,077	14,026	14,814	5,483	4,830	23,822	24,398	25,873	16,822	16,056
Mean expected present value of lifetime utility at age 16:											
No subsidy	185.6	156.5	162.5	154.8	142.9	133.9	189.6	208.6	198.4	182.0	174.2
Subsidy	197.4	165.2	171.6	166.2	143.0	134.1	222.9	231.3	223.5	196.1	185.0

Note: * The per capita cost of the subsidy is $2,247.

** The percent sickness up to and including age 20 without subsidy is 12.2%, while the percent sickness with subsidy is 11.3%.

1. Based on a simulation of 8,000 individuals.
2. Type 1: high ability in school and work; Type 2: high ability in school and low ability in work
   Type 3: low ability in school and high ability in work; Type 4: low ability in school and low ability in work
3. The skill endowments are drawn according to the population ratio of types.
4. The illness and duration are drawn from the initial health limitation distribution at age 16.

Table 2.13: Effect of a $778 Health Expenditure Subsidy for High School Students on Selected State Variables*

Characteristics	All	Sick up to and including Age 20**					Healthy up to and including Age 20				
		All Types	Type 1	Type 2	Type 3	Type 4	All Types	Type 1	Type 2	Type 3	Type 4
Average highest schooling years completed:											
No subsidy	13.39	12.17	12.35	12.72	10.12	10.24	13.56	13.64	14.03	12.0	12.26
Subsidy	13.92	13.20	13.25	13.64	11.76	12.03	13.99	13.89	14.43	13.18	13.31
Mean years in college											
No subsidy	1.85	0.73	0.65	1.12	0.05	0.07	2.01	2.19	2.34	0.43	0.61
Subsidy	2.29	1.68	1.72	2.13	0.34	0.50	2.35	2.22	2.79	1.61	1.86
Asset at age 30											
No subsidy	19,134	11,725	12,464	13,237	5,474	4,794	20,163	21,047	21,369	14,099	13,597
Subsidy	22,603	16,770	16,312	19,831	8,519	13,564	23,230	23,644	24,621	18,316	17,983
Mean expected present value of lifetime utility at age 16:											
No subsidy	195.4	156.5	162.5	154.8	142.9	133.9	200.8	208.6	198.4	182.0	174.2
Subsidy	218.5	186.6	192.4	187.8	159.7	164.9	221.7	228.0	220.4	204.7	198.6

Note: * The per capita cost of the subsidy is $2247, same amount as the per capita college tuition subsidy.

** The percent sickness up to and including age 20 without subsidy is 12.2%, while the percent sickness with subsidy is 9.7%.

1. Based on a simulation of 8,000 individuals.
2. Type 1: high ability in school and work; Type 2: high ability in school and low ability in work
   Type 3: low ability in school and high ability in work; Type 4: low ability in school and low ability in work
3. The skill endowments are drawn according to the population ratio of types.
4. The illness and duration are drawn from the initial health limitation distribution at age 16.

Chapter 3: Subjective Mortality Risk and Bequest Motivation

3.1 INTRODUCTION

It is known in the literature that a significant portion of household wealth is passed from one generation to another by bequest. According Kotlikoff and Summers (1981), 80% household-held capital was inherited. Gale and Scholz (1994) estimate that total bequests were $105 billion in the U.S. in 1986. Hurd and Smith (2002) find that the elderly anticipate leaving roughly 40% of their wealth in bequests. Kotlikoff (1988) claims that inherited wealth plays an important and perhaps dominant role in U.S. wealth accumulation. Thus it is conceivable that bequests may hold a key answer to the social security problem that baby boomers may face: they may eventually receive significant estates from their parents such that their dependence on social security may be reduced.

However, predicting whether a large portion of wealth will be passed from one generation to the next generation requires knowledge of the motives for bequests.[1] As pointed out in the literature (Kotlikoff 1988; Hurd 1989), a large amount of bequeathed wealth does not necessarily imply a substantial motive for bequests. Without a well-functioning annuity market, people will have to save against mortality risk, and the resulting bequests are involuntary. If most bequests are in fact involuntary or accidental, the value of the bequeathed wealth may decrease in the future as the annuity market

[1] Various incentives for bequest are offered in the literature. Some argue that bequests serve as incentives to younger generations to provide appropriate care for older generations (Cox 1987; Bernheim, Shleifer and Summers, 1985). Others argue that bequests are mainly motivated by altruism.

further develops.[2] In addition, it is also possible that people may change their perceptions of stock market risks after the recent crash of the market. In that case, more people may move into annuities, and the total amount of bequeathed wealth will decrease.[3]

There is no consensus in the literature on the significance of bequest motives. Some people (Bernheim 1987; Kotlikoff and Summers, 1988) argue that the bequest motive is important while others (Hurd 1989) claim that it is almost zero, and most bequests are accidental or involuntary.

It is well known that subjective expectations about future events are important factors to understand individual economic behaviors, such as saving, consumption and investment. However, few available information or data on individual subjective expectations limits the application of economic models to explain individual actions. Our main goal in this study is to investigate the empirical relevance of subjective survival rates for bequest motives. More specifically, we want to exam if and to what extent subjectively expected mortality risks are correlated with bequest motives in the presence of a borrowing constraint for single elders. We estimate a life cycle model with uncertain lifetime as developed by Yaari (1965) and Hurd (1989). Instead of applying the commonly used life tables to approximate individual survival expectations, we adopt the estimated individual subjective survival curves from Gan, Hurd and McFadden (2003, henceforth GHM).

Empirical estimates that are based on life-table survival curves are likely to be biased. For example, consider a typical utility function:

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$$

---

[2] Poterba (1997) documents that variable annuity premium payments increased by a factor of five during the period 1988-1993.

[3] The S&P 500 index peaked on August 2001 at 1517.7. Since then, it has dropped to 879.8 at the end of 2002.

where c_t is the consumption at time t, and γ is the risk aversion parameter. The first order condition in a common formulation (without a bequest motive) is:

$$\Delta \ln c_t \approx (r + \ln \beta + \Delta \ln s_t)/\gamma + f(X_t),$$

where X_t represents some socio-demographic and/or economic variables, r is the interest rate, and β is the time discount factor. s_t is the subjective survival probability at time t so that $-\Delta ln s_t$ is the mortality hazard rate. If s_t is not measured but it is correlated with X_t, we have a classic problem of endogeneity. If s_t is measured with error, the parameter estimate of γ will be biased.

A large panel dataset, the Asset and Health Dynamics among Oldest Old (AHEAD) collects data on people who were born between 1890 and 1923 and their spouses (regardless of age) including information on individuals' expectations of a wide range of future events. Respondents in the survey are asked about their expectations of chances to live to a certain age. Earlier work, such as Hurd and McGarry (1995, 2002) and GHM have looked at the subjective probabilities regarding survival rates. These papers have found that, on average, individual subjective survival probabilities are consistent with life tables and that they vary appropriately with known risk factors. Therefore, there is important information content in these responses on subjective survival probabilities.

However, the subjective survival probabilities suffer serious focal response problems: some individuals tend to give responses of 0.0 and 1.0. These focal responses cannot be directly used in analyzing life-cycle models where survival probabilities are required. To eliminate focal biases, GHM suggest a Bayesian update method. For each individual in the AHEAD data set, GHM estimate an "optimism" index. Compared to the life table survival probability, an individual may overestimate or underestimate his/her survival probability. The estimated "optimism" indices show significant individual heterogeneity, and can be applied to derive individuals' subjective survival probabilities without focal biases.

The rest of the study is organized as follows. In Section 2, we introduce a life-cycle model with bequests. Our emphasis is on how to estimate such a model. Section 3 presents the estimation results. In particular, Section 3.1 introduces the data that will be used in the study. Three key variables are used in the empirical variables: wealth, income and subjective survival probabilities. In Section 3.2, we present parameter estimates based on various estimation methods. Section 3.3 calculates the bequest incentives based on estimates from Section 3.2. In Section 3.4, we conduct out-of-sample predictions and simulate the consumption and wealth trajectories under various sets of parameter estimates. Finally, we summarize the results of this study and discuss several issues for future research in Section 4.

3.2 THE MODEL

Our starting point is the standard life-cycle model with bequest as in Yaari (1965) and Hurd (1989). Let the utility function of a retired individual be:

$$\sum_{t=0}^{N} \beta^t U(c_t) s_t + \sum_{t=0}^{N} \beta^t B(w_{t+1}) m_{t+1} \tag{3.1}$$

where s_t is the subjective probability that the individual will be alive at time *t*. m_{t+1} is the subjective mortality rate at time t + 1: $m_{t+1} = s_t - s_{t+1}$. The subjective maximal number of periods an individual can survive is *N*. The time discount factor is denoted as β. Consumption at time *t* is denoted as c_t, and the wealth at the beginning of time *t* is denoted as w_t. The first term in (3.1) is the present value of utility from consumption conditional on survival; and the second term in (3.1) is the present value of the utility from leaving a bequest of w_{t+1} conditional dying at $t + 1$. The utility from bequest, $B(w_{t+1})$, is increasing in w_{t+1}.

This model only applies to singles. The corresponding model for couples is much more complicated because it has to account for bequeathing by a couple to the next generation, and also for providing to a surviving spouse.[4]

As in Hurd (1989), we further assume a borrowing constraint such that bequeathable wealth cannot become negative. The constraint imposed on borrowing indicates that future Social Security benefits cannot be used as collateral for a consumption loan. This constraint arises from the fact that all heads of households in the sample are older than 70 years old in 1993 when the survey started, and in the U.S., Social Security benefits cannot be used as collateral. Such a constraint imposes important boundary condition in our analysis. Equation (2) lists the budget constraint at time t:

$$w_t = (1+r)w_{t-1} + A_{t-1} - c_{t-1} \geq 0, \tag{3.2}$$

where A_{t-1} is annuity income at time t-1.

It is typical in this literature to assume a constant risk aversion utility function $U(c_t) = c_t^{1-\gamma}/(1-\gamma)$. The income from annuities such as Social Security is assumed to be constant. The marginal utility of a bequest, denoted as α, is dependent on how many children the person has:

$$B_w \equiv \alpha \equiv \frac{\partial B}{\partial w} = 1_{\text{children}}\left(\alpha_0 + \alpha_1 * \text{No. of children}\right), \tag{3.3}$$

where 1_{children} is an indicator function. The assumption that the bequest motive exists only if the person has any children is important to identify the model. Otherwise, the identification may only come from the functional form assumptions.

[4] Estimating the couple's bequest motive is our next research objective.

The maximal age that a person may live, denoted as N, is obtained when the person's subjective survival rate $s_t < 1e\text{-}4$. Different agents have different maximum ages N since their subjective survival rates are different. Given the interest rate r, income A, and the parameter values of β, γ, and α, the paths of wealth are always contingent on the initial wealth w_0. However, the paths of consumption may not dependent on the initial wealth w_0. The solution to the optimization problem depends on whether the borrowing constraint is binding or not. The analysis of the solution of the discrete model is similar to that of the continuous model in Hurd (1989). Here we only state how to estimate the model.

Estimating the model requires at least two waves of wealth data for each individual. We use wealth data in wave 2 and wave 3 to estimate the model. The wave 4 wealth data is used for out-of-sample prediction.[5] The wealth level in wave 2 serves as the initial wealth w_0. We use backward induction to find the trajectories of the wealth and consumption. For a given set of parameter values β, γ, and α, we can obtain the trajectories of wealth $\{w_t^b, t = 1, \cdots, N+1\}$, where the superscript b indicates the value is calculated from backward induction. We then compare w_3^b at the trajectory with the observed wave 3 wealth w_3^*. We use the subscript 3 because in our data set the interval between the two waves of wealth is 3 years. The parameter set that minimizes the difference between w_3^b and w_3 are our estimates.

There are three types of consumptions paths corresponding to low, medium, and high wealth. We discuss these three different cases in the discrete model:

(1) In the first case, the bequest is strictly positive even if the individual survives to the greatest age possible: i.e., $w_{N+1} > 0$. Then the consumption trajectory satisfies:

5 There is strong evidence that wave 1 wealth data in AHEAD underestimate the stock ownership and hence the value of the stock wealth.

$$c_t^{-\gamma} s_t = \alpha \sum_{i=t}^{N} \beta^{i-t} (1+r)^{i-t} m_{i+1} \qquad (3.3a)$$

The trajectory in (3.3a) and actually initial wealth, w_0, satisfy

$$w_{t+1} = (1+r)^{t+1} w_0 + \sum_{i=0}^{t} (1+r)^{t-i} (A_i - c_i) > 0 \qquad (3.3b)$$

Equation (3.3a) shows that consumption trajectory is dependent of subjective survival rate but is independent of initial wealth w_0 if the wealth level at $N+1$ is strictly positive. This occurs because the marginal utility from consumption (left-hand-side) at time t has to equal to the present value of the marginal utility from bequest, which is assumed to be independent of wealth level. The wealth trajectory, w_t^b, can be calculated from the equation (3.3b), which shows that wealth trajectories vary according to the initial wealth w_0. Figure 1-1 shows the typical consumption and wealth trajectories. As illustrated, while wealth monotonically increases and consumption monotonically decreases, they may exhibit other patterns. The only requirement for this case is the wealth is strictly positive at any time in this person's life span.

The minimal level of initial wealth that corresponds to the consumption path (3.3a) is w_0^*, given by:

$$w_0^* = \sum_{i=0}^{N} (1+r)^{-i-1} (c_i - A) > 0 .$$

Any initial wealth larger than $w_0 > w_0^*$ will produce a consumption path $\{c^*\}$ as in (3.3a), and will lead to $w_{N+1} > 0$. Note that both N and w_0^* vary as individual subjective survival rate varies.

(2) In the second case, although the bequest is zero at the time of death, (w_{N+1}= 0), the borrowing constraint is not binding; that is, the wealth level is strictly positive for any $t<N+1$. The consumption path satisfies:

$$c_t^{-\gamma} s_t = \beta (1+r) c_{t+1}^{-\gamma} s_{t+1} + \alpha m_{t+1}, \text{ for } \quad t = 0,1,\ldots N-1 \qquad (3.4a)$$

$$w_{N+1} = (1+r)^{N+1} w_0 + \sum_{i=0}^{N} (1+r)^{N-i} \left(A_i - c_i\right) = 0 \tag{3.4b}$$

$$w_t > 0 \text{ , for } t = 1,2,\ldots N. \tag{3.4c}$$

Equation (3.4b) states that the consumption trajectory should lead to zero wealth level at time $N+1$: the person will leave no bequest should he or she live to the greatest age possible. Figure 3.1-3.2 illustrates one case where wealth reaches zero exactly at the maximum possible age. Consumption in Figure 3.1-3.2 first increases and then decreases as mortality risk becomes large. However, it is possible that consumption monotonically decreases if the time discount factor is small.

There will be a range of initial wealth and associated consumption paths that satisfy (3.4a), (3.4b) and (3.4c). The intuition for this result will be discussed when we provide estimation algorithm (Step 2 in the algorithm. See Appendix B). Let w_0^* be the largest of these values so that any value of w_0 larger than w_0^* leads to $w_{N+1} > 0$ and the consumption path will be independent of w_0. Let $\hat{w}_0$ be the smallest of those values so that any smaller value of initial wealth causes the wealth to reach 0 before $N+1$. Let $\{\hat{c}\}$ and $\{\hat{w}\}$ be the individual's consumption and wealth trajectories associated with $\hat{w}_0$, and $\{c^*\}$ and $\{w^*\}$ be the individual's consumption and wealth trajectories associated with w_0^*. Therefore, in the case of medium wealth, the consumption trajectory must lie between $\{\hat{c}\}$ and $\{c^*\}$, and the wealth trajectory must lie between $\{\hat{w}\}$ and $\{w^*\}$.

(3) Lastly, we consider the case that the borrowing constraint is binding. Let T be the time when bequeathable wealth is exhausted. The consumption path is found from the solutions to four equations, (3.5a)-(3.5d):

$$c_t = A \text{ for } t = T,\cdots,N\text{ ,} \tag{3.5a}$$

$$c_t^{-\gamma} s_t = \beta(1+r) c_{t+1}^{-\gamma} s_{t+1} + \alpha\, m_{t+1}\text{ , for } t = 0,1,\ldots T-2\text{ ,} \tag{3.5b}$$

$$w_T = (1+r)^T w_0 + \sum_{i=0}^{T-1} (1+r)^{T-1-i} \left(A_i - c_i\right) = 0\text{ .} \tag{3.5c}$$

$w_t > 0$, for $t = 1, 2, \ldots T-1$. (3.5d)

In this case consumption and wealth will eventually decline. Figure 3.3 illustrates consumption and wealth trajectories in this case.

Each individual in our sample has a different subjective survival curve. Therefore, every individual's critical value of wealth is different. We search to find out his/her critical wealth value, and then calculate his/her consumption and wealth trajectories. Our objective is to find a set parameter values that minimize the difference between the predicted second wave wealth, w_3^b, and the observed second wave wealth, w_3. We consider two different objective functions: mean square loss function and the absolute value loss function.

$$\min_{\alpha,\beta,\gamma} \sum_i \left(w_{3i} - w_{3i}^b\right)^2 \tag{3.6a}$$

$$\min_{\alpha,\beta,\gamma} \left\{ \sum_i \left| w_{3i} - w_{3i}^b \right| \right\} \tag{3.6b}$$

The mean square loss function in (3.6a) is the one used in Hurd (1989). The absolute value loss function in (3.6b) corresponds to median regression. The advantage for median regression over the mean regression is that median regression is robust to outliers.

We briefly discuss how to estimate the covariance matrix. Let the parameter set be denoted as $\delta = (\gamma, \beta, \alpha)'$, and let the covariance matrix be Ω. It is straightforward to obtain the covariance matrix for estimates based (3.6a). The covariance matrix from median regression in (3.6b) is given by:

$$\Omega = \frac{1}{4 f_u^2(0)} \left(E\left[\left(\frac{\partial w_3^b}{\partial \delta} \right) \left(\frac{\partial w_3^b}{\partial \delta} \right)' \right] \right)^{-1}, \tag{3.7}$$

where $f_u(0)$ is the density of the error term u evaluated at 0. The error term u is defined as $u = w_3 - w_3^b$. Empirically, we first conduct a non-parametric kernel regression, and then evaluate the obtained density function at 0 to get $f_u(0)$. The expectation part can be calculated by sample average. Since no explicit solutions exist for the derivative $\partial w_3^b / \partial \delta$, numerical derivatives are used in the calculation.

3.3 DATA AND ESTIMATION RESULTS

3.3.1 Data

Our data set consists of the second, third and fourth waves of the AHEAD sample. We do not employ wave 1 data because there is good evidence that the first wave of AHEAD underreported asset holdings. To select our sample, we use the following sample selection criteria: (1) Because the model in this study applies only to singles, our sample only includes people who are alive and who are singles in both wave 2 and wave 3. (2) Total wealth or non-housing wealth is non-negative in wave 2 and wave 3. (3) Responses to the survival probability question in wave 2 are valid. When total wealth is used as one of the selection criterion, the number of valid observations is 1,903. When we consider non-housing wealth, the number of observations decreases to 1,752. Among these valid observations in wave 1 and wave 2, only 1,460 of them are still valid in wave 3.

Three key variables are used in this study: household wealth, income, and individual subjective survival curves. We now discuss these three variables in detail.

(1) The Wealth and Income Data

The AHEAD survey is a panel survey of older Americans. The wave 1 survey of AHEAD was conducted in 1993. The initial sample of AHEAD includes a sample of people who were 70 years old or more in 1993 (and their spouses regardless of ages). The wave 2 survey was conducted in 1995, and waves 3 and wave 4 were conducted in 1998 and 2000, respectively. The AHEAD data set provides more than 10 categories of wealth data. It is well-known in the literature that often a large portion of people do not provide valid responses on wealth questions (Juster and Smith, 1997; Chand and Gan, 2003). AHEAD uses a sequence of questions to bracket a wealth item. Although this technique is very successful in reducing non-response rates, it requires serious effort to impute the wealth values. Chand and Gan (2002) discuss various imputation methods. The imputed wealth data used in this study are obtained from Adams *et al* (2003) who impute three waves of wealth and two waves of income. In Table 3.1, we list summary statistics of the total wealth and the wealth net of housing wealth. For each wave of wealth, we list the mean, median, variance, minimum and maximum values. From Table 3.1, mean wealth decreases slightly between wave 2 and wave 3 but decreases significantly between wave 3 and wave 4. Specifically, between wave 2 and wave 3, the mean total wealth decreases 4.5% while the non-housing wealth decreases by 2.5%. Between wave 3 and wave 4, the mean total wealth declines 18% and the non-housing wealth declines 30%. The pattern for the median wealth is different from the mean wealth. Between wave 2 and wave 3, the median wealth decrease 14% and 15% for total wealth and non-housing wealth. However, between wave 3 and wave 4, there is a slightly increase for the median total wealth with rate 5.8%. The decreasing rate for non-housing wealth between wave 3 and 4 is 6.2%.

As Table 3.1 indicates, the median wealth is less than half of the mean wealth, reflecting the positive skewness that exists in the asset distribution. More specifically, the median is respectively 35%, 32% and 48% of the mean for the wave 2, 3 and 4 total wealth, and respectively 20%, 14%, and 19% of the mean for the wave 2, 3, and 4 non-housing wealth.

In Table 3.2, we list age, the number of children and income. The average age of respondents in the second wave is 79 years of old. Although heads of households in our sample have to be at least 72 years in wave 2, their spouses who may be younger are also included in the sample. The number of people in our sample who are younger than 72 years old is 46 (2.63% of the sample). Among all the people in our sample, 80.2% have children. The average number of children in our sample is 2.55. One household has 16 children. Second wave income is used as a measure of people's annuity income. The mean income level is $18,107 with a large standard deviation of $22,873.

(2) Individual Subjective Survival Probability

In this study, for each individual, we construct two survival curves: the life-table survival curve and the subjective survival curve. The life-table survival curve is directly obtained from the life table. The subjective survival curve is obtained from GHM. Here we briefly describe the subjective survival curve. One innovation in two recent surveys (Health and Retirement Study and AHEAD) is that they include questions about respondent's subjective probabilities about events in the future. In particular, each respondent is asked about his/her perceived probability of surviving to a target age that is between 10 and 15 years in the future. Although Hurd and McGarry (1995, 2002) show that on average these subjective probabilities are generally consistent with life tables, at the individual level, they suffer a serious problem. In all age groups, a substantial fraction of respondents give responses of 0.0 and 1.0. These responses cannot represent the respondents' true probabilities. GHM develop a model to recover each individual's "true" subjective probability.

Given the same age and sex, different people may have very different subjective survival probabilities. Some of the difference may relate to the health and wealth situations of individuals, some may simply be reflect personality. For each individual in their data set (AHEAD), GHM estimate an “optimism” index. Compared to the life table survival probability, an individual may overestimate or underestimate his/her survival probability. The estimated “optimism” index in GHM shows that significant individual heterogeneity exist in the AHEAD population. In a simple life cycle model, GHM show that ignoring individual heterogeneities may result in bias estimates. In this study, we apply both the subjective survival probability developed in GHM and the life table survival probability.

Four different “optimism” indices were estimated in GHM, representing four different specifications. In this study, we use the “unconstrained hazard-scaling” index.[6] In particular, let the current age of individual *i* be *a*. His subjective survival probability to age $a+t$ is given by:

$$s_{ia}(t) = \exp\left(-\int_0^t \lambda_{ia}(a+r)dr\right),$$

where $\lambda_{ia}(a+t)$ is the hazard function at age $a+t$. Further, let the individual’s life table hazard be $\lambda_{i0}(a+t)$. The “unconstrained hazard-scaling” in GHM assumes that: $\lambda_{ia}(a+t)=\psi_i\lambda_{i0}(a+t)$ where ψ_i is the individual’s optimism index. If $\psi_i>1$, this individual is said to be “pessimistic”; if $\psi_i<1$, then this person is “optimistic”. Table 3.2 has the summary statistics of this index estimated fro responses in wave 2.

The mean and median of ψ_i are .659 and .663, respectively. People in this sample are on average more optimistic about their survival probabilities than the life table implies. A more optimistic person may save more than a life-table person would do. If we use an observed sequence of wealth to estimate our model, the estimates based on subjective survival curves should indicate a lower time discount factor and/or lower bequest motive than the estimates based on life tables.

---

6 We select this index because it has the best predictive power of actual survival experience among all four indices.

3.3.2 Estimation Results

Our main results exclude housing wealth. In principle, at the extreme of very high transaction costs, it is difficult to change the consumption level of housing. Therefore, holding of housing wealth would simply reflect initial conditions and differences between the rate of housing appreciation and the general inflation rate. Excluding housing wealth from bequeathable wealth would give a better idea of the change in desired wealth holdings than would be found from including housing wealth.[7]

In Table 3.3, we report the estimates of our model by using non-housing wealth and by assuming a fixed interest rate r = 0.04. We will test the robustness of our estimates later by using different interest rates. In Panel (A) of Table 3.3, we apply median regression to estimate the model using both subjective and life-table survival curves. Although the marginal utility of bequests is estimated to be almost zero in both cases, other parameter estimates vary significantly. Using life-table survival curve yields a higher time discount rate than using subjective survival curves. This is expected because people subjectively overestimate their survival probabilities. They behave accordingly by saving more to prepare for a longer lifespan, rather than valuing future consumption more than current-period consumption as implied by the estimates based life-table survival curves.

Panel (B) in Table 3.3 lists the estimates when the mean regression method is used. The marginal utilities of bequest in this panel are much larger than those estimated in Panel (A), which imply strong bequest motives. Another observation in Panel (B) is

---

[7] For completeness, however, we also estimated the model over total wealth, which includes housing asset. The results over total wealth actually are very close to those over non-housing wealth. For example, the estimates over total wealth and subjective survival rates for parameters risk-averse coefficient γ, time discount factor β, and bequest motive parameter α_0, and α_1 are 0.9088, 0.9468, 4.9759e-7, 1.0272e-6, respectively.

that the time discount factor is estimated to be significantly larger than 1, indicating that people value future consumption more than current consumption. Between the two sets of the estimates from mean regressions, the time discount factor is higher when the life table survival curve is used.

It is important to note that in a life-cycle model of time-varying survival probabilities, a time discount factor that is larger than 1 does not imply necessarily non-stationary growth in either consumption or wealth. Kocherlakota (1990) shows that it is possible that people still prefers current consumption to future one even with β>1, as long as the output or income grows at a rate that is sufficiently high. Kocherlakota's discussion is based on an infinitely lived representative agent. In our model, individual agent has constant income levels. From equation (3.1), even with β>1, the rate of consumption growth will turn negative at the time when the hazard rate $-\Delta \ln s_t$ is large enough.

The reason to have such an unusual time discount factor is that non-housing wealth during the sample period declined by only 2.5%. Apparently because of no significant difference in bequest motives between those who have children and those who do not have children, the marginal utility of bequest is almost always zero. Given the constant interest rate at .04, matching such a small decrease in wealth requires the individual to have tremendous incentives to save. This large saving incentive has to come from a large time discount factor. One major drawback, we suspect, is the interest rate we use: the return to capital investment may not be at 4% during our sample period. However, how to formally incorporate varying interest rate requires a model of portfolio choice. The dataset does not have enough information to estimate such a model.

In summary, mean regression yields very different parameter estimates from median regression. More specifically, mean regression suggests very large desired bequests while the median regression implies almost zero bequest motives. In addition, life table mortality risk yields a slightly larger bequest motives than subjective mortality risk.

In Table 3.4, we list results from median regressions with varying interest rates. The risk-averse parameters and the time discount factor are very close to the reference value when interest rate changes from .02 to .06. Within this range of interest rates the marginal utility of bequests is very small.

In the following section, we will try to understand the economic significance of the bequest motives by some simulation exercises.

3.3.3 Bequest Simulations

Among the three parameters we estimate, it is relatively easy to understand the economic significance of the risk-averse parameter γ and the time discount factor β. To understand the effect of γ and β on bequests, consider a familiar consumption growth equation in the absence of the bequest motive of equation: $\Delta \ln c_t \approx (r + \ln \beta + \Delta \ln s_t)/\gamma$. Given the survival rate s_t and the risk-averse parameter γ, a larger β will increase algebraically the slope of the consumption path and because of the lifetime budget constraint, initial consumption will have to be reduced. Thus more wealth will be held and so bequests will increase. Although the effect of time discount factor β on bequests is clear, the effect of the risk-averse parameter on bequests is ambiguous. When the consumption path is decreasing a larger γ will increase algebraically the slope of the consumption path causing more wealth to be held and increasing bequests. When the consumption path is increasing a larger γ will flatten the consumption path causing initial consumption to be higher but later consumption to be lower. Therefore, the total effect on bequests or wealth holdings for γ is ambiguous. It is important to note that a change in bequests because of a change in either γ or β is a change in accidental bequest.

A non-accidental bequest is measured by the marginal utility of bequest, α. The larger the value of α, the larger is the bequest motive.

Two methods measure the economic significance of marginal utility of bequest, α:

$$\sum (1+r)^{-t}\left[\hat{w}_t(\hat{\alpha}) - \hat{w}_t(0)\right] m_t \tag{3.8a}$$

$$\sum \left[\hat{w}_t(\hat{\alpha}) - \hat{w}_t(0)\right] s_t \tag{3.8b}$$

where $\hat{\alpha} \equiv 1_{\text{children}}\left(\hat{\alpha}_0 + \hat{\alpha}_1 \cdot \text{No of children}\right)$. In (3.8a) and (3.8b), $\hat{w}_t(\hat{\alpha})$ is the optimal wealth trajectory given initial wealth and the estimated values of parameters. The term $\hat{w}_t(0)$ is defined in the similar way except that the marginal utility of bequests is zero. Equation (3.8a) and (3.8b) represent two different ways to understand the effect of bequests. In (3.8a), we calculate the present value of bequests. In (3.8b), we calculate the population difference in wealth holdings with and without a bequest motive. In another words, (3.8b) represents the effect of a bequest motive on the population wealth holdings. In Table 3.5, we calculate the effect of a bequest motive for a particular individual: a male at age 79 whose initial wealth is \$35,000 and whose income is \$12,000. The individual has two children. The optimistic index of this individual is 0.6594, and thereby the life expectancy is 109 years old.

The results in Table 3.5 are presented in three different panels, grouped by their estimation methods. In the first three rows, (R1)-(R4), we let the marginal utility of bequests vary. In particular, row (R1) corresponds to a bequest motive estimated from (A1) in Table 3.3 where subjective mortality risk is used. We let time discount factor vary in rows (R5)-(R7), and let the risk averse parameter vary in rows (R8)-(R10). The marginal utility of bequest parameter has significant impact on the level of desired bequest and on the difference in wealth holdings. In rows (R1)-(R4) where the risk averse parameter (γ) and the time discount factor (β) are estimated using the median regression, the desired bequest rises from almost zero to \$125,278 and the difference in wealth holding increases from \$1 to \$1,082,618 when the marginal utility of bequests increases from 2.47E-06 to 1. The effect of varying the marginal utility of bequests on desired bequests and on the difference in wealth holdings is very large. As the marginal utility of bequests is 1, the consumption path is rather flatten decreasing from \$1,211 at age 79 to \$1,013 age 109, which implies that the agent's 90% - 95% annuity of \$12,000 are turned into bequests or wealth holdings. In contrast, as the marginal utility of bequests is the value from median regression with subjective mortality risk, the consumption path is rather deep dramatically dropping from \$21,766 at age 79 to annuity level \$12,000 at age 86.

In rows (R5)-(R7), we allow the time discount factor vary while keeping risk averse parameter constant. The marginal utility of bequest is constant at 0.001. In this case, desired bequests increases from \$2.58 to \$1,408 when the time discount factor increases from 0.7 to 1.3. The result that a larger time discount factor is related to a higher desired bequest is consistent to the prior discussion. Finally, in rows (R8)-(R10), we consider the effect of risk averse parameter γ. A larger γ implies a more risk averse agent. When γ increases from 0.5 to 2.0, the desired bequest increases from \$5.80 to \$518.5.

In summary, simulation results show that a higher marginal bequest motive, larger time discount factor, and larger risk averse parameter all increase the level of desired bequests significantly. A modest increase in either of the three variables may lead to a very large increase in desired bequests and in differences in wealth holdings.

3.3.4 Consumption/Wealth Trajectory and Out-of-Sample Predictions

A typical way to evaluate parameter estimates from different methods is to conduct out-of-sample predictions. We used wealth data in wave 2 and wave 3 to obtain parameter estimates. We will now use the estimated parameters to predict the wealth values in wave 4, and compare the predicted wealth to observed wealth in wave 4. Table 3.6 has the comparison results. Each column in Table 3.6 reports various sums of errors based on a given set of parameter estimates. The column number, A1, A2, B1, or B2, corresponds to the estimates listed in Panel A and Panel B in Table 3.3. These estimates differ in their estimation method and their survival probabilities. The out-of-sample calculation is based on the same survival probability as the parameter estimates are. For example, if the set of parameters is obtained based on subjective survival probability, the out-of-sample calculation is also based on the subjective survival probability.

Parameter estimates in Column (A1) and (A2) are from median regressions while Column (B1) and (B2) are from mean regressions. Not surprisingly, (A1) and (A2) have smaller absolute errors and smaller mean square errors than (B1) and (B2), regardless of error types. Furthermore, (A1) and (A2) have a lower sum of absolute errors for low wealth people and a larger sum of absolute errors for high wealth people than (B1) and (B2). This is expected because mean square regressions tend to fit high-wealth observations better because the large wealth values are magnified by the square operation.

Results in Table 3.6 can also be used to evaluate the advantage of using subjective survival probability instead of life-table survival probability. When the median regressions are used, parameter estimates based on subjective survival probability (A1) produce lower sums of mean square errors and lower sum of absolute errors in out-of-sample prediction of wealth than estimates based on life-table survival curves. In particular, the mean square errors and the absolute errors from subjective survival curves are 42% and 5% less than the corresponding errors from life-table survival curves.

The second and the third panel in Table 3.6 report comparison results based on predicted mean and predicted median. Although predicted means using both survival curves are lower than the observed mean at wave 4, the mean ($87,033) from subjective survival curves is much closer to the observed mean ($118,112) than the mean ($71,413) from life-table survival curves. Further, we divide the sample into four quartiles according to the wealth level at wave 3, and compare the predicted and observed means in each quartile. In the fourth panel in Table 3.6, using subjective survival curves produces better predictions than using life-table survival curves in all four quartiles. At the first quartile, the predicted mean using subjective survival curves is $8.6 while the predicted mean using life table is $2,385. The observed means at wave 4 is $-1,548. At the second quartile, the predicted mean from subjective survival curves is $7,947, which is much closer to the observed mean ($9,091) than the predicted mean from life table ($2,385). Similar patterns are observed for the third and fourth quartiles.

When the mean regression method is used, parameter estimates based on subjective survival curves do not have a significant advantage in predicting fourth wave wealth comparing to ones based on life-table survival curves. Based on these results, we conclude that median regression is better than mean regression, and subjective survival probability better describes individual saving and bequest decisions than the life-table survival probability.

Finally, to better understand how people's consumption and wealth vary, we apply estimates from Table 3.3 to simulate a hypothetical person's consumption and wealth trajectories in Figure 3.2. The hypothetical person we consider is: single male at age 79 with an optimistic index of .6594. He has two children. His initial wealth and income are assumed at the median values in Table 3.2. In addition, the parameter set for Figure 3.2 is obtained from the median regression in Table 3.3. His consumption level is highest when he starts at age 79, and decreases until he reaches age 85.His wealth decreases and reaches zero at age 85. Above age 85, the person's wealth keeps reaches zero and his consumption equals to his annuity income at $12,000. If the person dies before age 85, he leaves some bequest. However, such bequest is accidental since his bequest motive is essentially zero. In all these cases, since the person values future utility lower than current utility, his consumption level peaks at the first year and then decreases until it reaches his annuity income level.

3.4 CONCLUSION

This study investigates if and to what extent bequest motives exists for a sample of single elderly people. Our main goal in this study is to estimate a classical life-cycle model with bequests, as in Yaari (1965) and Hurd (1989) for elderly with individual-specific subjective survival curves. In almost any life-cycle models, individual mortality risk is an important factor that affects people's decisions. Previous literature assumes the individual mortality risk is the same as the life-table mortality risk, ignoring the apparent individual heterogeneity in their mortality risk. This assumption may cause biases in parameter estimates. This study applies the individual subjective survival probability model developed in an earlier study (GHM). Their subjective survival probabilities have significant variations across individuals, and can better predict actual survival experience than life tables. We find that the estimation results from mean regressions differ significantly from median regression results. Most importantly, mean regression yields very large desired bequests while the median regression implies almost zero bequest motives. In addition, we find that life table mortality risk yields a little bit larger bequest motives than subjective mortality risk.

Table 3.1: Summary Statistics of Wealth

(Being alive and single in the 2nd and 3rd waves; wealth is not negative; not missing subjective survival question; in 1995 dollars)

	wave 2		wave 3		wave 4	
	total wealth	non-housing wealth	Total Wealth	non-housing wealth	total wealth	non-housing wealth
Mean	221,728	173,042	211,760	168,634	174,428	118,112
median	78,500	35,000	67,190	23,364	70,746	22,500
std dev	1,416,500	1,446,572	1,299,766	1,253,508	404,712	317,598
Minimum	0	0	0	0	-52,632	-157,895
Maximum	43,325,000	43,225,000	36,794,393	31,186,916	8,368,421	5,679,825
No. of obs	1903	1752	1903	1752	1460	1460

Table 3.2: Summary Statistics

	Mean	std dev	Median	min	max
Age of respondents in 1995	79	5.21	78	63	92
Income in wave 2					
Sample of 1903 observations	17,764	22,146	12,000	468	466,000
Sample of 1752 observations	18,107	22,873	12,000	468	466,000
Percentage who have children	80.2%				
Number of children	2.5514	2.3028	2	0	16
Survival probabilities					
optimism index (ψ)	0.6594	0.1176	0.6631	0.4385	1.0906
subjective 3-year survival prob	0.8911	0.0509	0.9026	0.6225	0.9893
life-table 3-year survival prob	0.8347	0.0844	0.8592	0.4175	0.9790
no. of observations in the sample	1752				

Table 3.3: Estimation Results:

(Marginal Utility of Bequest = 1_{child}*($\alpha_0 + \alpha_1$ * No. of kids), interest rate = .04, non-housing wealth)

	estimation method	subjective or life table	risk averse parameter (γ)	time discount rate (β)	marginal utility of bequest (α_0)	marginal utility of bequest (α_1)
A1	median	subjective	0.9855	0.9420	3.8067e-7	1.0431e-6
			(0.0519)	(0.0028)	(8.957e-5)	(4.6931e-5)
A2	Median	life table	0.7403	1.0045	7.6864e-4	2.1185e-5
			(0.1275)	(0.0044)	(8.601e-4)	(1.7597e-4)
B1	Mean	subjective	0.7870	1.0546	1.0008	1.0022
			(1.544)	(0.8767)	(0.1525)	(0.925)
B2	Mean	life table	0.7634	1.0763	0.9986	0.8941
			(1.295)	(0.6890)	(0.2316)	(0.7546)

Table 3.4: Robust Test with Median Regression Results

(varying interest rates, subjective survival rate, non-housing wealth)

interest rate used (r)	risk averse parameter (γ)	time discount rate (β)	marginal utility of bequest (α_0)	marginal utility of bequest (α_1)
0.02	0.8933	1.0151	1.7789e-5	1.8797e-6
	(0.1960)	(0.0061)	(3.3e-3)	(7.9283e-4)
0.03	0.8053	1.0049	7.2723e-6	3.57e-6
	(0.1797)	(0.0050)	(2.8102e-3)	(8.4822e-4)
0.04	0.9855	0.9420	3.8067e-7	1.0431e-6
	(0.0519)	(0.0028)	(8.957e-5)	(4.6931e-5)
0.05	0.9783	0.94	9.7635e-46	1.3841e-50
	(0.2420)	(0.0163)	(2.6350e-020)	(4.8609e-020)
0.06	0.9007	0.9293	9.1176e-48	1.468e-44
	(0.0289)	(0.0029)	(3.1365e-21)	(6.1125e-21)

Table 3.5: Economic Significance of Marginal Utility of Bequest

(For a hypothetical person: male, age 79, 2 kids, optimist index = 0.6594, initial wealth = \$35,000, income = \$12,000)

rows	Risk averse parameter (γ)	time discount rate (β)	Marginal utility of bequest ($\alpha_0 + 2*\alpha_1$)	Desired bequest	Difference in wealth holdings
R1	0.9855	0.942	2.4669e-6	\$0.05	\$1.17
R2	0.9855	0.942	.001	\$21.12	\$477.22
R3	0.9855	0.942	.1	\$32,316	\$514,790
R4	0.9855	0.942	1	\$125,278	\$1,082,618
R5	0.9855	0.70	.001	\$2.59	\$57.26
R6	0.9855	1.00	.001	\$80.48	\$1,434
R7	0.9855	1.20	.001	\$1,408	\$18,238
R8	0.	0.9420	.001	\$5.80	\$116.7
R9	1.5	0.9420	.001	\$129.5	\$2,413
R10	2	0.9420	.001	\$518.5	\$9,463

Table 3.6: Results from Out-of-Sample Predictions

Models	med reg (subjective) (A1)	med reg (life table) (A2)	mean reg (subjective) (B1)	mean reg (life table) (B2)
		Error Comparison		
mean square error	6.5230e8	1.1248e9	2.6798e9	2.7650e9
absolute error	1.5489e5	1.6440e5	2.6789e5	2.6744e5
		Mean Comparison		
predicted mean	87,033	70,719	249,913	247,281
observed mean		118,112		
		Median Comparison		
predicted median	14,795	71,413	96,540	95,1780
observed median		22,500		
		Comparison by Quartile[1]		
The first quartile				
predicted mean	8.6	617.7	33,221	33,791
observed mean		-1,548		
The second quartile				
predicted mean	7,946.7	2,385	74,516	74,004
observed mean		9,091		
The third quartile				
predicted mean	36,853	23,305	147,202	145,170
observed mean		53,905		
The fourth quartile				
predicted mean	3.0189e5	2.5647e5	7.4251e5	7.3481e5
observed mean		3.5151e5		

[1] The sample is divided in quartiles according the observed 3rd wave wealth.

Figure 3.1 Illustration of the Positive Bequest Case

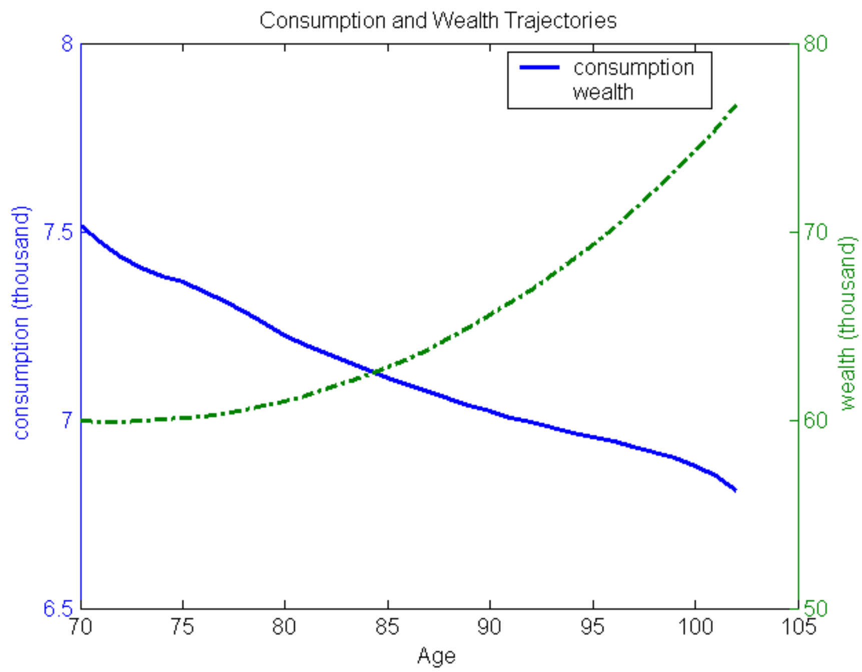


Figure 3.2 Illustration of the Zero Bequest Case

(No Borrowing Constraint Binding)

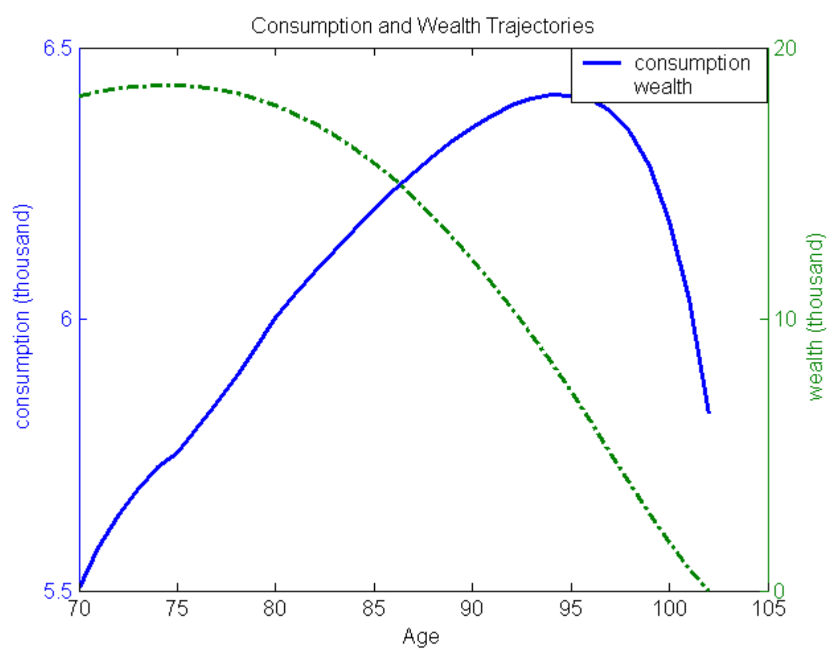
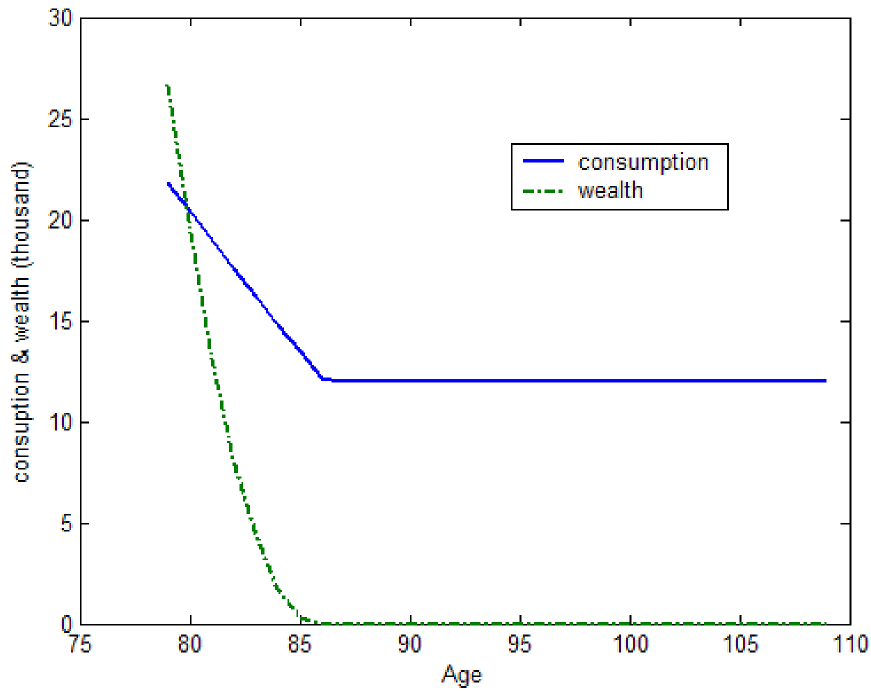


Figure 3.3: Consumption and Wealth Trajectories at Median Wealth Level [a]



[a] a hypothetical person: male, age 79, 2 kids, optimistic index .6594, initial wealth \$35,000, income \$12,000;
risk averse $\gamma = 0.9855$, time discount $\beta = 0.9420$, bequest motive: $\alpha_0 = 3.8067\text{e-}7$, $\alpha_1 = 1.0431\text{e-}6$; desired bequest is \$0.05, and difference in wealth holdings is \$1.17.

Appendices:

APPENDIX A: ESTIMATION METHOD FOR CHAPTER 2

A.1. Forms of Descriptive Statistical Models

The criteria for choosing an appropriate descriptive statistical model are computational tractability and statistical efficiency in which it can provide a good description of the data. The linear probability models, as suggested by Keane and Smith (2003), fit the criteria precisely.

In the present context, the linear probability models are constructed according to the observed data. Thus, the structures of the models differ because of the differences in the missing and unobserved data across the time periods. The descriptive statistical model at time *t* is given as

$$y_{it} = x_{it}\eta_t + \upsilon_t , \qquad \upsilon_t \sim iidN(0, \Sigma_t) \tag{A.1}$$

where y_t is a vector of independent variables, x_t is the vector of regressors, $\theta_t = (\eta_t, \Sigma_t)$ is the set of parameters to be estimated.

The construction of the descriptive models is described by time periods as follows. For convenience, I use the notations for the observed data to describe the model.

(1) $t = 1$, i.e., age 16

The regressors include a constant term, schooling years, indicator of sickness, and duration of sickness:

$$x_1 = (1, sch_1, D_1, sl_1). \tag{A.2}$$

Note that because of the data limitation in calculating the effective schooling years *edu*, I use the observed schooling years *sch.* The indicator for the success in the school is chosen as an independent variable.

The set of dependent variables consists of the choice of working, wage, choice of schooling, indicators of passing the grade, and indictors of sickness at period two. The dependent variables are allowed to differ in their sizes. If for some individuals, one or more dependent variables were missing or unobserved, then the corresponding dependent variables are accordingly missing from these samples. For example, the transcript data were missing or unobserved for individual *i* (unobserved transcript may occur because he was in middle school or college during the time of survey), then the dependent variable of the indicator of the passing of a grade will not be included for this individual. The set of the dependent variables for the observed data is

$$y_1 = \left(d_{1,1}^1, w_1, d_{2,1}^1, pass(high\, school), D_2\right). \tag{A.3}$$

The simulated data consists of the same individuals as in observed data, except that the simulated discrete variables are replaced by the smooth functions discussed in the Section 2. That is to say that the number of linear regression equations for simulated data and observed data is equal.

(2) $1 < t < 6$, i.e. from age 17 to 20

For t = 3, 4 or 5, the regressors include a constant term, schooling years, working experiences, choice of job participation, choice of school attendance, indicator of sickness, and duration of sickness:

$$x_t = \left(1, sch_t, ep_t, d_{1,t}^1, d_{2,t}^1, D_t, sl_t\right), \tag{A.4.1}$$

The independent variables for t = 2 are different from those for t = 3, 4 or 5, in which working experience was not included because at this period ep_2 is equal to $d_{1,1}^1$ (remember that the initial working experience is set at zero):

$$x_2 = \left(1, sch_2, d_{1,1}^1, d_{2,1}^1, D_2, sl_2\right), \tag{A.4.2}$$

The dependent variables are:

$$y_t = \left(d_{1,t}^1, w_t, d_{2,t}^1, pass(high\,school), D_t\right). \tag{A.5}$$

Similar to the case of t = 1, if some observed variables were missing or unobserved, the corresponding dependents were also missing.

(3) t = 6, i.e., age 21. At this age, some agents start to have asset data and some do not. The set of the independent variables are the same as in (A.4.1). The set of dependent variables are

$$y_6 = \left(d_{1,6}^1, w_6, d_{2,6}^1, pass(high\,school), D_7, A_6\right), \tag{A.6}$$

(4) 6 < t < 16, i.e., from age 22 to 30

Two descriptive statistical models are set apart by the asset data. Both of the models have the same set of dependent variables:

$$y_t = \left(d_{1,t}^1, w_t, d_{2,t}^1, D_{t+1}, A_t\right). \tag{A.7}$$

Note that the indicator for passing the grade is not included in (A.7) because of the convenient assumption that individuals should have finished their high school by age 22. Actually, in my data only 5 samples were in high school over 21 years old.

The first model includes all the individuals whose assets at $t-1$ were missing or unobserved. In contrast, the second model includes all the individuals whose asset at $t-1$ were observed. The set of the independent variables for the first model is the same as in (A.4.1), while for the second one it is

$$x_t = \left(1, sch_t, ep_t, d_{1,t-1}^1, d_{2,t-1}^1, D_t, sl_t, A_{t-1}\right) \tag{A.8}$$

(5) t = 16, i.e. age 32

The descriptive statistical models are similar to the case of $6 < t < 16$, in which the models are distinguished by whether the assets at period 15 were observed. The set of independent variables for the first model is

$$x_{16} = \left(1, sch_{16}, ep_{16}, d^1_{1,15}, d^1_{2,15}, D_{16}, sl_{16}\right) \tag{A.9.1}$$

and for the second model is

$$x_{16} = \left(1, sch_{16}, ep_{16}, d^1_{1,15}, d^1_{2,15}, D_{16}, sl_{16}, A_{15}\right) \tag{A.9.2}$$

Because the sample does not contain the information for health at $t = 17$, the set of dependent variables is

$$y_{16} = \left(d^1_{1,16}, w_{16}, d^1_{2,16}, A_{16}\right) \tag{A.10}$$

A.2. Two-Step Approach

For the sake of computational tractability, I use the two-step approach as proposed by Keane and Smith (2003) to estimate the parameters of the structural model. The idea of the first step is to obtain a consistent estimate $\hat{\psi}_1$ of the structural parameters by solving the optimization problem (2.21) in Section 2. In the first step, the number of simulated data sets F is set to 1, which substantially reduce the computation time. In addition, a relatively large value for the smoothing parameter λ is chosen ($\lambda = 0.05$) to ensure the objective function is smooth.

In the second step, to reduce bias I choose λ equal to 0.003 and F equal to 100. According to Proposition 2 in Keane and Smith (2003),

$$\hat{\psi}_2 = \hat{\psi}_1 - \left(\hat{J}' L_{\Theta\Theta}\left(y; z, \tilde{\tilde{\Theta}}(\hat{\psi}_1)\right)\hat{J}\right)^{-1} \hat{J}' L_{\Theta\Theta}\left(y; z, \tilde{\tilde{\Theta}}(\hat{\psi}_1)\right)$$

is a consistent and asymptotically normal estimate of ψ_0, where $L_{\Theta\Theta}$ is the Hessian of the likelihood function associated with the descriptive model and $\hat{J}$ is an estimate of the Jacobian of the binding function $\mathrm{H}(\hat{\psi}_1)$.

APPENDIX B: ALGORITHM TO FIND THE OPTIMAL CONSUMPTION AND WEALTH PATH IN CHAPTER 3

Step 1: Check the high wealth case, in which a strictly positive bequest is left at the maximum age of life, i.e., $w_{N+1} > 0$.

(1) From equation (3.3a), we calculate the consumption trajectory $\{c_t^b, t = 0, \cdots N\}$.

(2) Substitute the trajectory of consumption $\{c_t^b, t = 0, \cdots N\}$ into Equation (3.3b) to get the wealth trajectory $\{w_t^b, t = 1, \cdots, N+1\}$.

(3) If for all $t \in \{1, 2, \ldots, N\}$, $w_t^b \geq 0$ and $w_{N+1}^{\;b} > 0$, then report w_3^b and go to next observation; else go to *Step 2*.

Step 2: Check the medium wealth case, in which the wealth at the end of maximum age of life is zero, i.e., $w_{N+1} = 0$, and at all other time periods $t \leq N$, $w_t > 0$. We use backward induction to get the consumption and wealth trajectories.

(1) From (3.4a), c_t ($t = 0, \ldots, N$ -1) is a function of c_N by recursive iteration: $c_t = c_t(c_N)$. Substitute the trajectory of consumption $\{c_t(c_N), t = 0, \ldots, N\text{ -1}\}$ into Equation (3.4b) such that wealth level in (3.4b) now is only a function of c_N. In particular, we have:

$$w_{N+1}(c_N, w_0) = 0 \tag{B1}$$

Given observed w_0, we can solve (B1) to get c_N, denoted as c_N^b. Given c_N^b, we can apply (3.4a) to iteratively find out $\{c_t^b, t = 0, \cdots, N-1\}$. However, if we do not know w_0, we will have many values of c_N and w_0 such that (B1) are satisfied. Among them, the higher bound w_0^* is the maximum of w_0 such that (B1) is satisfied and $c_t > 0$ for all $t < N+1$; the lower bound $\hat{w}_0$ is the smallest w_0 such that (B1) is satisfied and $c_t > 0$ for all $t < N+1$.

(2) If for all $t \in \{0, 1, \ldots, N\}$, $c_t^b > 0$, then calculate the wealth trajectory $\{w_t^b, t = 1, \cdots, N\}$ from Equation (3.2); else go to *Step 3*.

(3) If for all $t \in \{1, 2, \ldots, N\}$, $w_t^b > 0$, then report w_3^b and go to next observation; else go to *Step 3*.

Step 3: Check the low wealth case, in which the wealth reaches zero at a time period $T \leq N$. We search all over the possible T from the backward. The method is similar to *Step 2*.

(1) Let $T = N$. From (3.5b), c_t ($t = 0, \ldots, T-2$) is a function of c_{T-1} by recursive iteration: $c_t = c_t(c_{T-1})$. Substitute the trajectory of consumption $\{c_t(c_{T-1}), t = 0, \ldots, T-2\}$ into Equation (3.5c) such that (3.5c) now is only a function of c_{T-1}. Solve the equation: $w_T = 0$ to get c_{T-1}, denoted as c_{T-1}^b. We can get the consumption trajectory $\{c_t^b, t = 0, \cdots, N\}$ by applying (3.5b) with given c_{T-1}^b.

(2) If for all $t \in \{0, 1, \ldots, T-1\}$, $c_t^b > 0$, then calculate the wealth trajectory $\{w_t^b, t = 1, \cdots, T-1\}$ from Equation (3.2); else let $T = T-1$, and repeat (1) - (2).

(3) If for all $t \in \{1, 2, \ldots, T-1\}$, $w_t^b > 0$, then break from the cycle, report w_3^b and go to next observation; else let $T = T-1$, and repeat (1) - (3).

References

Adams, Peter, Michael Hurd, Daniel McFadden, Angela Merrill, and Tiago Ribeiro (2003). "Healthy, Wealthy, and Wise? Tests for Direct Paths between Health and Socioeconomic Status," Journal of Econometrics, Vol 112 (2003): 3-56.

Altonji, Joseph G. and Rebecca M. Blank, 1999. "Race and gender in the labor market," in O. Ashenfelter and D. Card (eds), Handbook of Labor Economics, Volume 3c, Elsevier Science Press. (1999): 3144–3259.

Arrow, K. (1971). "Essays of Theory of Risk Bearing." Chicago: Markham Publishing Company.

Backus, David K., Patrick J. Kehoe, and Finn E. Kydland, 1994. "Dynamics of the trade balance and the terms of trade: the J-curve?" American Economic Review, 84(1) (March 1994): 84–103.

Bernheim, B. Douglas, Andrei Shleifer, and Lawrence Summers (1985). "The Strategic Bequest Motive," Journal of Political Economy, Vol 93 (1985): 1045-1076.

Bils, Mark and Peter Klenow, 2000. "Does schooling cause growth?" American Economic Review, 90(6) (December 2000): 1160–1183.

Cameron, Steven and James J. Heckman (1993), "The Nonequivalence of High School Equivalents," Journal of Labor Economics 11: 1-47.

Cameron, Steven and James J. Heckman (1998), "The Dynamics of Educational Attainment for Blacks, Hispanics, and Whites," Photocopied. Chicago: University of Chicago.

Chand, Harish, and Li Gan (2003). "The Effect of Bracketing in Wealth Estimation," Review of Income and Wealth, 49(2), (June 2003): 273-87.

Chand, Harish, and Li Gan (2002). "Wealth Item Non-response and Imputation Methods," mimeo, Department of Economics, University of Texas, Austin (2002).

Cox, Donald (1987). "Motives for Private Income Transfer," Journal of Political Economy, Vol 95 (1987): 508-546

Currie, J. and Hyson, R. (1999), "Is the Impact of Health Shocks Cushioned by socioeconomic Status? The Case of Low Birth Weight", American Economic Review 89: 245-50.

Durlauf, Steven, 2002. “Groups, social influences and inequality: a membership theory perspective on poverty traps.” Working paper, Department of Economics, University of Wisconsin.

Ehrlich, Isaac and F.T. Lui, 1991. “Intergenerational trade, longevity, and economic growth.” American Economic Review, 99(6) (October 1991): 1029–1059

Fernandez, Raquel and Richard Rogerson, 1996. “Income distribution, communities, and the quality of public education.” Quarterly Journal of Economics, 111: 135–164.

Fries, James F., 1980. “Aging, natural death, and the compression of morbidity.” New England Journal of Medicine, 303(3): 130–136.

Fuchs, V.R. (1982): “Time Preferences and Health: An Exploratory Study,” in: V.R. Fuchs, ed., Economic Aspects of Health (University of Chicago Press for the National Bureau of Economic Research, Chicago): 93-120.

Gale, Williams G., and John Karl Scholz (1994). “Intergenerational Transfers and the Accumulation of Wealth,” Journal of Economic Perspectives, Vol 8 No 4 (1994): 145-160.

Gan, Li, Michael Hurd, and Daniel McFadden (2003). “Individual Subjective Survival Curves,” NBER Working Paper Series, #9480 (February 2003).

Gourieroux, C., A. Monfort, and E. Renault (1993), “Indirect Inference,” Journal of Applied Econometrics 8, S85-S118.

Grossman, Michael, 1972a. “On the concept of health capital and the demand for health.” Journal of Political Economy, 80: 223–255.

Grossman, Michael, 1972b. The Demand for Health: A Theoretical and Empirical Investigation. New York: Columbia University Press, for the National Bureau of Economic Research.

Grossman, Michael (1975), “The Correlation between Health and Schooling,” in: N.E. Terleckyj, ed., Household Production and Consumption (Columbia University Press for the National Bureau of Economic Research, New York) 147-211.

Grossman, Michael (1999), “The Human Capital Model of the Demand for Health,” National Bureau of Economic Research Working Paper: 7078.

Grossman, Michael, 2000. “The human capital model of the demand for health,” in J. Newhouse and A. Culyer (eds), Handbook of Health Economics, Amsterdam: North Holland.

Grossman, Michael and R. Kaestner (1997), "Effects of Education on Health," in J.R. Berhman and N. Stacey Eds. The Social Benefits of Education. University of Michigan Press, Ann Arbor, 1997.

Hurd, Michael D (1989). "Mortality Risks and Bequests," Econometrica, Vol 57, n4 (July 1989): 779-813.

Hurd, Michael D, Daniel McFadden and Li Gan (1998). "Subjective Survival Curves and Life Cycle Behavior." In David Wise, ed, Inquiries of Economics of Aging, Chicago: The University of Chicago Press (1998), 259-305.

Hurd, Michael D. and Kathleen McGarry (2002). "The Predictive Validity of Subjective Probabilities of Survival." The Economic Journal, Vol 112, No 482 (October 2002).

Hurd, Michael, and James Smith. "Expected Bequest and Their Distribution," NBER Working Paper Series, #9142 (September 2002).

Jones, Charles, 2002. "Why have health expenditures as a share of GDP risen so much?" Working Paper w9325, National Bureau of Economic Research

Kalemli-Ozcan, Sebnem, Harl E. Ryder, and David N. Weil, 2000. "Mortality decline, human capital investment, and economic growth." Journal of Development Economics, 62 (2000): 1–23.

Kamien, M.I. and N.L. Schwartz, 1991. Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management, (2d ed), New York: North-Holland.

Keane, Michael P., and K. Wolpin (1994), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence", Review of Economics and Statistics 76: 648-72.

Keane, Michael P. and K. Wolpin (1997): "The Career Decisions of Young Men," Journal of Political Economy 105: 473-522.

Keane, Michael P., and Anthony A. Smith (2003), "Generalized Indirect Inference for Discrete Choice Models," working paper.

Kenkel, Donald S. (1991), "Health Behavior, Health Knowledge, and Schooling," Journal of Political Economy 99: 287-305

Kenkel, Donald S., 2000. "Prevention," in J. Newhouse and A. Culyer (eds), Handbook of Health Economics, Amsterdam: North Holland.

Kotlikoff, L., and L. Summers (1981). "The role of intergenerational transfers in aggregate capital accumulation," Journal of Political Economy, 89, 706-732.

Juster, F. T., and James Smith (1997). "Improving the Quality of Economic Data: Lessons from HRS and AHEAD," Journal of American Statistical Association, Vol 92, No 440 (1997): 1268-78.

Levy, Haim (1994), "Absolute and Relative Risk Aversion: An Experimental Study", Journal of Risk and Uncertainty, Vol. 8, No. 3 (May 1994), pp.289-307

Levy, Helen and David Meltzer (2001), "What Do We Really Know About Whether Health Insurance Affects Health?" University of Chicago Harris School of Public Policy, Unpublished Working Paper.

Livi-Bacci, M., 1997. A Concise History of World Population (2nd ed), Oxford: Blackwell. Translated by Carl Ipsen.

Massey, Douglas and Nancy Denton, 1989. "Hypersegregation in U.S. metropolitan areas: Black and Hispanic segregation along five dimensions." Demography, 26 (August 1989): 371–391.

Manski, Charles, 1993. "Identification of endogenous social effects: the reflection problem." Review of Economic Studies, 60(3), 531–542.

Matthews, R.C.O., C.H. Feinstein, and J.C. Odling-Smee, 1982. British Economic Growth 1856–1973, Stanford, California: Stanford University Press.

Maxwell, Nan, 1994. "The effect on black–white wage differences on differences in quantity and quality of education." Industrial and Labor Relations Review, 47(2): 249–264.

National Center for Health Statistics, 1987. U.S. decennial life tables for 1979–1981, 1(1), Hyattsville, Maryland.

Neal, Derek A. and William R. Johnson, 1996. "The role of pre-market factors in black-white wage differences." Journal of Political Economy, 104(6): 869–895.

Newhouse, Joseph P (1993), Free for All? Lessons from the RAND Health Insurance Experiment. Cambridge and London: Harvard University Press

O'Neill, June, 1990. "The role of human capital in earnings differences between black and white men." Journal of Economic Perspectives, 4 (Fall): 25–45.

Perri, Timothy J. (1984), "Health Status and Schooling Decisions of Young Men," Economics of Education Review.

Preston, Samuel H., 1980. "Causes and consequences of mortality declines in less developed countries during the twentieth century," in Richard S. Easterlin (ed), Population and Economic Change in Developing Countries. National Bureau of Economic Research (Chicago: The University of Chicago Press, 1980): 289–341.

Preston, Samuel H., 1996. "Population studies of mortality." Population Studies, 50(3), (November 1996): 525–536.

Rosenzweig, Mark R. and T. Paul Schultz, 1983. "Estimating a household production function: heterogeneity, the demand for health inputs, and their effects on birth weight." Journal of Political Economy, 91(5) (October): 723–746.

Rosenzweig, M. R. and Schultz, T. P. (1991), "Education and Household Production of Child Health", in Proceedings of the American Statistical Association (Social Statistics Section) (Washington, DC: American Statistical Association).

Roux, Ana V.D., Sharon Stein Merkin, Donna Arnett, Lloyd Chambless, Mark Massing, Javier Nieto, Paul Sorlie, Moyses Szklo, Herman A. Tyroler, and Robert L. Watson, 2001. "Neighborhood of residence and incidence of coronary heart disease." The New England Journal of Medicine, 345(2): 99–106.

Sen A.K., 1999. Development as Freedom, New York: Knopf.

Soares, Rodrigo Resis, 2002. "Life expectancy, education attainment, and fertility choice." Working paper, Department of Economics, University of Maryland.

Winkleby, M.A. and C. Cubbin, 2003. "Influence of individual and neighborhood socioeconomic status on mortality among black, Mexican-American, and white women and men in the United States." Journal of Epidemiology and Community Health, 57 (June).

Winship, Christopher and Sanders Korenman, 1997. "Does staying in school make you smarter?" in Intelligence, Genes, and Success, Devlin et al. (eds), Spring-Verlag.

Yaari, Menahem (1965). "Uncertain Lifetime, Life Insurance and the Theory of the Consumer," Review of Economic Studies, Vol 32 (1965): 137-150.

Vita

Guan Gong was born in Yanglinwei, Xiantao in Hubei province in the People's Republic of China. He is the son of Chunyun Gong and Zaihui Tian. He received the Bachelor of Science degree in Mathematics in 1992 and the Master of Science degree in Applied Mathematics in 1995, from the Wuhan University. He worked as an assistant professor in Institute of Advanced Economics Study at Wuhan University and Guanghua Management School of Pekinng University from 1996-2000. He started the doctoral degree program at the University of Texas at Austin in August 2000.

Permanent address: 1025 Babu Street, Apt. 701, Wuhan, Hubei 430072, P.R.China

This dissertation was typed by Guan Gong.